
Precision calculations in the Higgs sector within and beyond the Standard Model

Matteo Capozzi



München 2020

Precision calculations in the Higgs sector within and beyond the Standard Model

Matteo Capozzi

Dissertation
an der Fakultät für Physik
der Ludwig-Maximilians-Universität
München

vorgelegt von
Matteo Capozzi
aus Colleferro

München, den 14/02/2020

Erstgutachter: Prof. Dr. Gerhard Buchalla
Zweitgutachter: Prof. Dr. Wolfgang Hollik
Tag der mündlichen Prüfung: 20/03/2020

This thesis is based on the author’s work conducted at the Max Planck Institute for Physics (Werner-Heisenberg-Institut) in Munich from February 2017 until February 2020 . Parts of this work have already been published or are in preparation for publication:

Refereed Research Papers

- M. Capozzi, G. Heinrich, S. P. Jones, M. Kerner, *Probing anomalous Higgs boson couplings in Higgs plus jet production at NLO QCD with full m_t -dependence*, [in preparation].
- M. Capozzi, G. Heinrich, *Exploring anomalous couplings in Higgs boson pair production through shape analysis*, *JHEP* **03** (2020), 091[1908.08923].
- G. Buchalla, M. Capozzi, A. Celis, G. Heinrich, M. Kerner, L. Scyboz, *Higgs boson pair production in non-linear Effective Field Theory with full m_t -dependence at NLO QCD*, *JHEP* **09** (2018) 057 [1806.05162].

Conference Proceedings

- M. Capozzi, G. Heinrich, *Shape analysis in Higgs boson pair production*, in *14th International Symposium on Radiative Corrections: Application of Quantum Field Theory to Phenomenology (RADCOR 2019) Avignon, France, September 8-13, 2019*, 2019 1912.07039.
 - M. Cepeda et al., *Higgs Physics at the HL-LHC and HE-LHC, HL/HE WG2 group report*, 1902.00134.
- G. Buchalla, M. Capozzi, A. Celis, G. Heinrich, M. Kerner, L. Scyboz, *Higher order and top mass effects in Higgs boson pair production beyond the Standard Model*, *PoS* **LL2018** (2018) 083 1807.11839.

Zusammenfassung

Die Entdeckung einer mit dem Higgs-Boson des Standard Modells (SM) konsistenten Resonanz am Large Hadron Collider (LHC) im Jahre 2012 markierte das Ende der Suche nach dem letzten vom SM vorhergesagten Teilchen. Dadurch wurde der Mechanismus der elektroschwachen spontanen Symmetriebrechung experimentell bestätigt. Zurzeit existiert kein Hinweis, der dem SM widerspricht und die Präsenz "neuer Physik" attestieren würde. Neue Teilchen könnten theoretisch im TeV-Bereich auftreten: Üblicherweise erfolgt die Parametrisierung von neuen, über das Standardmodell hinausführenden (BSM) Effekten durch eine effektive Feldtheorie (EFT). Dabei werden solche BSM Effekte in anomalen Kopplungen codiert.

Momentan ist der Higgs-Sektor wenig untersucht worden, und lässt einigermaßen Abweichungen von SM-Vorhersagen zu. Wir präsentieren Resultate für zwei Prozesse, die grosse Bedeutung am LHC haben: Die Produktion von Higgs-Boson-Paaren (hh), sowie die Produktion eines Higgs-Bosons assoziiert mit einem Jet ($h + 1j$). In beiden Fällen werden Wirkungsquerschnitte und differentielle Verteilungen bis zur nächstführenden Ordnung (NLO) in QCD gezeigt. Diese werden mit voller Top-Massenabhängigkeit im Rahmen einer im Higgs-Sektor nicht-linear realisierten EFT erzeugt. Mithilfe dieser Rechnungen wird der Raum der anomalen Kopplungen untersucht, sowie deren Effekte auf physikalische Observablen.

Die NLO QCD Korrekturen zur Produktion von Higgs-Boson-Paaren, die durch fünf anomale Kopplungen charakterisiert ist, werden präsentiert. Wir analysieren die Auswirkung der NLO Korrekturen auf Verteilungen der invarianten Masse des Di-Higgs-Systems, sowie des transversalen Impulses eines der Higgs-Bosonen. Wir extrahieren eine Parametrisierung des inklusiven Produktionswirkungsquerschnitts und der m_{hh} -Verteilung als Funktion der anomalen Higgs-Kopplungen samt NLO Korrekturen.

Ausserdem präsentieren wir eine Klassifizierung von möglichen Profilen, die in der Verteilung der invarianten Masse m_{hh} vorkommen. Hierfür verwenden wir zwei Ansätze: Eine Analyse, die auf prädefinierten Profiltypen basiert, und eine Klassifizierung, die durch unüberwachtes Lernen erzielt wird. Wir beobachten im Fall von Higgs + Jet, dass, im Vergleich zur konventionellen Profilanalyse, unsere auf unüberwachtem Lernen basierende Methode eine detaillierte Studie der Effekte erlaubt, die von den anomalen Higgskopplungen erzeugt werden.

Schließlich präsentieren wir die Berechnung von NLO QCD Korrekturen zur Higgs-Bosons assoziiert mit einem Jet Produktion. Die Berechnung ist in einem Code für Monte-Carlo Event-Erzeugung implementiert, welcher öffentlich zugänglich sein wird. Die Streuamplitude des Prozesses ist eine Funktion von zwei anomalen Kopplungen. Bestimmte Konfigurationen der Kopplungen erzeugen einen ähnlichen Wirkungsquer-

Zusammenfassung

schnitt. Wir analysieren zwei dieser Konfigurationen und zeigen, wie sie unterschieden werden können.

Abstract

The discovery of a boson consistent with the Standard Model (SM) Higgs boson in 2012 at the Large Hadron Collider (LHC) finally closed the hunting for the last unobserved SM particle and experimentally confirmed the electroweak symmetry breaking mechanism. Nevertheless there are still open questions suggesting extensions of the SM. At the moment there is no evidence contradicting the SM and proving the presence of New Physics. New Particles could likely show up in the TeV range, if it is the case a good choice is to parametrise the Beyond Standard Model (BSM) effects in an Effective Field Theory framework. The New Physics effects are encoded in anomalous couplings. At the moment the Higgs sector is least explored and allows deviations from SM predictions, giving hints to extend this sector.

We present the study of two processes of particular interest at the LHC: Higgs boson pair production and Higgs plus one jet production. In both cases we present the next-to-leading order (NLO) QCD cross section and differential distributions with full top mass dependence within the framework of a non-linearly realised Effective Field Theory in the Higgs sector. In these calculations we explore the anomalous couplings parameter space probing the effects of these parameters on the observables.

Concerning the Higgs boson pair production in the gluon fusion channel, characterized by five anomalous couplings, we present a calculation of the NLO QCD corrections where we analyse the impact of modifications of the anomalous couplings on the distributions in the Higgs boson pair invariant mass and the transverse momentum of one of the Higgs bosons. We provide an analytical parametrisation for the total cross-section and the m_{hh} distribution as a function of the anomalous Higgs couplings that includes NLO corrections.

We further present our shapes classification of Higgs boson pair invariant mass distributions m_{hh} . We use two approaches: an analysis based on predefined shape types and a classification into shape clusters based on unsupervised learning. We find that our method based on unsupervised learning allows a more detailed study of the impact of anomalous couplings on the m_{hh} shape compared to more conventional approaches to a shape analysis. Higgs jet.

At last we present the calculation of NLO QCD corrections for the Higgs plus one jet production. The computation is implemented in a Monte Carlo event generator code which will be publicly available. The scattering amplitude of the process is function of two anomalous couplings. Certain configurations of the couplings produce total cross sections which are almost the same. We study two of these configurations, and we show how to disentangle them.

Contents

Zusammenfassung	vii
Abstract	ix
Contents	xi
1 Introduction	1
2 The Standard Model	5
2.1 $SU(3)_C$	5
2.2 $SU(2)_L \times U(1)_Y$	7
2.3 Symmetry breaking mechanism	8
2.4 The Higgs boson self couplings	10
3 Higher order perturbation theory	13
3.1 Factorization theorem	13
3.2 Quantum Chromodynamics (QCD) perturbative expansion	14
3.3 Differential cross section distributions	15
3.4 Regularization and renormalisation	16
3.4.1 Dimensional regularization	17
3.4.2 Renormalisation	17
3.4.3 Infrared divergencies	18
4 Effective Field Theories formalism	21
4.1 Linear EFT	22
4.2 Non-linear EFT	22
5 Machine Learning tools	25
5.1 Perceptron and Fully Connected Network	25
5.2 Autoencoders	27
5.3 Unsupervised Learning through K-Means	29
6 Higgs boson pair production	31
6.1 Higgs boson pair production state of art	32
6.2 Non-Linear EFT for di-Higgs production	33
6.3 Calculation of the NLO QCD corrections	34
6.3.1 Virtual corrections	36
6.3.2 Real radiation	37

6.3.3	Parametrisation of the total cross section	38
6.3.4	Validation of the calculation	38
6.4	Phenomenological results	39
6.4.1	NLO cross sections and heat maps	40
6.4.2	Cross sections and distributions at several benchmark points . .	42
6.4.2.1	Total cross sections	42
6.4.2.2	m_{hh} and $p_{T,h}$ distributions	43
6.4.2.3	Discussion of the benchmark points	49
6.5	Conclusions	50
7	Higgs boson pair production and m_{hh} distribution shape analysis	57
7.1	Classification through predefined shape types	58
7.1.1	Definition of shape types	58
7.1.2	Classification of m_{hh} distributions	60
7.1.3	Classification of $p_{T,h}$ distributions	63
7.2	Classification and clustering by unsupervised learning	65
7.2.1	Unsupervised learning procedure	65
7.2.2	Parameter space underlying the clusters	66
7.3	Conclusions	71
8	Higgs plus one jet production	77
8.1	Higgs boson production state of art	78
8.2	Non-Linear EFT for Higgs plus one jet production	79
8.3	NLO QCD corrections	80
8.3.1	Leading order	80
8.3.2	Virtual corrections	81
8.3.3	Real radiation	83
8.4	Phenomenological results	83
8.5	Conclusions	88
9	Conclusion and Outlook	89
	List of Figures	93
	List of Tables	97
	Bibliography	99
	Acknowledgments	119

1 Introduction

The Standard Model (of particle physics) (SM) is the theory describing electroweak and strong interactions. SM parameters have been measured with excellent precision during the last decades. In 2012 at the Large Hadron Collider (LHC) two experiments, ATLAS and CMS, discovered a boson consistent with the SM Higgs boson [1, 2] and in 2015 its mass has been measured with high precision $m_H = 125.10 \pm 0.14$ GeV [3]. The Higgs discovery was the final consecration of the SM framework developed at the end of the 1970s. Nevertheless there are hints suggesting that the SM is not a complete theory such as neutrinos oscillation, dark matter and strong CP problem. In the last years several theories tried to extend the SM and to embed it in a more general framework. These theories are referred to as beyond Standard Model (BSM) theories. All of them involved new particles but at the moment there is no evidence of them. The LHC runs or other experiments have not found yet proof of new particles indeed. By the way recently a discrepancy with the SM predictions has been observed in the B-meson decay [4] which if confirmed could lead to the lepton flavor universality (LFU) violation. The absence of New Physics particles modified the strategy of the LHC, experiments are focusing more on precision measurements of the SM parameters than searches for new resonances. At date one of the LHC main goal is to explore the Higgs sector. Some properties of the Higgs boson have been measured with amazing precision, but others as the couplings to fermions and Higgs boson self couplings still allow deviations from SM predictions. Fig. 1.1 shows the total cross section for the Higgs boson production in association with a top quark-antiquark pair $pp \rightarrow t\bar{t}H$, where the top Yukawa coupling can be observed directly, as function of the invariant mass. The measured total cross section is 670 ± 90 (stat.) $^{+110}_{-100}$ (syst.) fb [5], which is consistent with the SM prediction of 507^{+35}_{-50} fb but it leaves room for possible new Physics contributions.

Another parameter the experimental community is trying to constrain is the triple Higgs self coupling, directly related to the Higgs potential but the cross sections of the processes where this parameter shows up are very small, leading to a loose constrain on it. We discuss this point more in detail in Sec. 2.4.

The actual status of experimental measurements shows that there is no evidence of New Physics but there is still space for deviations from SM in the Higgs sector which if firmly confirmed would be a sign of New Physics. The situation can be seen as new particles are hiding at energies higher of the ones reachable at the LHC, but the presence of these particles influences the scattering cross sections and so the experimental readouts. The New Physics could hide in TeV or higher scale, and an Effective Field Theory (EFT) framework can be the tool to parametrise New Physics effects without including the unknown new particles. Within an EFT framework their effects are represented by anomalous couplings whose magnitude has to be defined. The values

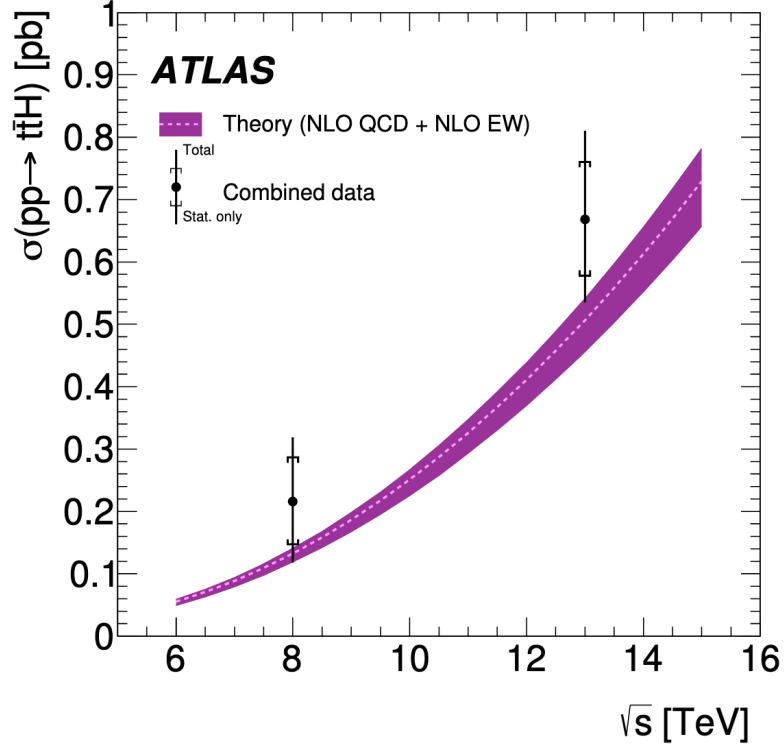


Figure 1.1: $pp \rightarrow t\bar{t}H$ total cross sections in proton-proton collisions at centre-mass energy of 8 and 13 TeV. The black dots represent the experimental data including uncertainties. The SM prediction as function of the energy \sqrt{s} is represented by the dashed line, where the band enveloping this line is the theoretical uncertainty [5].

they can assume are limited by theoretical and experimental constraints defining ranges where these couplings can be tuned and varied in order to study the impact of the deviations with respect to the SM predictions arising from the EFT parametrisation. If enough statistics are present the anomalous couplings can be fitted against experimental data.

Calculations within an EFT framework can help to explore the anomalous couplings space and show the possible deviations arising from New Physics effects, but to improve the sensitivity to them, next-to leading order and higher-order QCD corrections are crucial. The Higgs boson pair production in gluon fusion is an excellent process to study within an EFT framework, where the Higgs boson self coupling shows up at leading order. This process is very promising to find out whether the Higgs boson self-coupling is Standard-Model-like, but its cross section is very small therefore it is very hard to observe at the current energies. Within an EFT framework, the Feynman amplitude of the process depends on five anomalous couplings. The couplings can

change drastically the value of the total cross section or yield a degeneracy with the SM prediction. The differential distributions instead, change in magnitude and shape in response of the anomalous couplings modification. The shape analysis is a good opportunity to understand how each coupling affects the distributions. To achieve this task machine learning approaches come out to be very useful since they allow to select certain features minimizing the human bias.

There are other processes of interest in the Higgs sector which can be observed experimentally, such as single Higgs production, inclusive or with jets. The high statistic of single Higgs events allow to match theoretical predictions to experimental observations. Concerning the Higgs plus one jet production within an effective field theory framework, the process at LO depends on two anomalous couplings: the rescaling of the top Yukawa and the gluon-gluon Higgs effective coupling. This process is a window to observe deviation due to New Physics effect analyzing its total cross section and differential distributions and compare these observables to the experimental observations.

This thesis is organised as follows. The theoretical formulation and experimental confirmation of the SM is reviewed in Chapter 2. In Chapter 3 we present the techniques needed to carry out the theoretical predictions: the factorization theorem for hadron-hadron collisions, the QCD perturbative expansion and the treatment of the divergencies arising in next-to-leading order calculations. We present the EFT formalism in Chapter 4 giving an overview of the two possible realisations: linear and non-linear. In Chapter 5 we introduce machine learning techniques, reviewing the applications to physics problems and giving a general explanation of the algorithms we used in our calculations. In Chapter 6 we present a calculation of the NLO QCD corrections to Higgs boson pair production within the framework of a non-linearly realised Effective Field Theory in the Higgs sector. We analyze how the NLO corrections affect distributions in the Higgs boson pair invariant mass and the transverse momentum of one of the Higgs bosons. In Chapter 7 we investigate further the Higgs boson pair invariant mass distributions m_{hh} , we present our analysis where we classify distinct classes of shapes and visualise how they are related to the underlying coupling parameter space. We use two approaches: an analysis based on predefined shape types and a classification into shape clusters based on unsupervised learning. We show our calculation of Higgs plus one jet production at NLO QCD within the framework of a non-linearly realised EFT in Chapter 8. We analyze the impact of the modified top Yukawa coupling and the effective coupling on the total cross section and distribution in the transverse momentum of the Higgs boson.

2 The Standard Model

The SM is the theoretical framework describing the fundamental particles and their interactions, through the three fundamental forces: electromagnetism, weak interactions and strong interactions. The SM is a gauge invariant and renormalisable quantum field theory based on the symmetry group

$$SU(3)_C \times SU(2)_L \times U(1)_Y. \quad (2.1)$$

$SU(3)_C$ is the quantum chromodynamics (QCD) group of strong interactions while $SU(2)_L \times U(1)_Y$ is the chiral group of the electroweak interactions. The quantum numbers of this group are weak isospin I_3 and hypercharge Y for $SU(2)_L$ and $U(1)_Y$ respectively. The predictive power of SM has been extensively tested computing cross sections and differential distributions of a large variety of processes. The predictions yielded resulted always in agreement with the experiments and at date there are no significant deviation contradicting the SM. Because of this reliability it is widely accepted as the correct framework to describe the three fundamental forces and the microscopic world of matter.

2.1 $SU(3)_C$

$SU(3)_C$ is the QCD group, the theory describing the strong interactions among quarks and gluons. The group has eight generators T^a forming the Lie algebra:

$$[T^a, T^b] = if^{abc}T^c \quad (2.2)$$

where f^{abc} , the structure constants of $SU(3)$, are represented by an anti-symmetric tensor. The generators T^a in the fundamental and anti-fundamental representation are 3×3 matrices while in the adjoint representation are 8×8 matrices.

In 1954 Yang and Mills [6] formulated a non-Abelian Lagrangian symmetric under $SU(N)$. Its generalization for $N = 3$, summed over the quarks flavour, gives the Lagrangian specifying the QCD

$$\begin{aligned} \mathcal{L}_{QCD} &= \sum_q \bar{q}(i\gamma^\mu D_\mu)q - \frac{1}{4} \sum_{a=1}^8 G_{\mu\nu}^a G_a^{\mu\nu} \\ D_\mu &= \partial_\mu - ig_s T^a A_\mu^a. \end{aligned} \quad (2.3)$$

2 The Standard Model

q , transforming under the fundamental representation, is a triplet of quarks fields, containing the three colored quarks:

$$q = \begin{pmatrix} q_r \\ q_g \\ q_b \end{pmatrix}, \quad (2.4)$$

D_μ is the covariant derivative and g_s is the strong coupling. The tensor $G_{\mu\nu}^a$ can be expressed as

$$G_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + g_s f^{abc} A_\mu^b A_\nu^c \quad (2.5)$$

where A_μ^a , with $a = 1, \dots, 8$, is the gluon vectorial field, which appears as the adjoint representation of $SU(3)_C$. Loosely speaking \mathcal{L}_{QCD} specifies a theory where the colored quarks, fermions of spin $\frac{1}{2}$, interacting each others with the mediation of the gluons. The vector bosons, the gluons, are spin 0 particle carrying eight possible color charge and in contrast to the photons they can also interact among themselves.

The Lagrangian of eq. (2.3) does not contain mass terms for the quarks. We will see in sec. 2.3 how these particles obtain their masses.

The up (u), down (d) and strange (s), quarks have been introduced for the first time by Gell-Mann and Zweig [7, 8] as the fundamental components of hadrons: baryons are composed of three quarks and mesons of one quark and one anti-quark.

Within the GIM mechanism [9], which describes the flavour-changing neutral currents, the charm (c) quark has been predicted. Then with the extension of the Cabibbo matrix [10] into the Cabibbo-Kobayashi-Maskawa (CKM) matrix [11], which explains the CP violation and the down-type quarks eigenstates mixing to give the mass eigenstates, the top(t) and bottom (b) quarks have been introduced. The QCD particles content is composed of six quarks, grouped in three generations, and the gluon. All the QCD particles have been observed by experiments [12–19]. Table 2.1 show the quarks properties and quantum numbers, all the values are taken from Ref. [3].

Particle	Mass	J	B	Q	I ₃	Generation
up	$2.16^{+0.49}_{-0.26}$ MeV	$\frac{1}{2}$	$\frac{1}{3}$	$+\frac{2}{3}$	$+\frac{1}{2}$	First
down	$4.67^{+0.48}_{-0.17}$ MeV	$\frac{1}{2}$	$\frac{1}{3}$	$-\frac{1}{3}$	$-\frac{1}{2}$	First
charm	1.27 ± 0.02 GeV	$\frac{1}{2}$	$\frac{1}{3}$	$+\frac{2}{3}$	$+\frac{1}{2}$	Second
strange	93^{+11}_{-5} MeV	$\frac{1}{2}$	$\frac{1}{3}$	$-\frac{1}{3}$	$-\frac{1}{2}$	Second
top	172.9 ± 0.4 GeV	$\frac{1}{2}$	$\frac{1}{3}$	$+\frac{2}{3}$	$+\frac{1}{2}$	Third
bottom	$4.18^{+0.03}_{-0.02}$ GeV	$\frac{1}{2}$	$\frac{1}{3}$	$-\frac{1}{3}$	$-\frac{1}{2}$	Third

Table 2.1: Table reassuming the quarks properties. J is the total angular momenta, B the baryon number, Q the electric charge and I₃ the isospin.

The QCD coupling g_s^2 , as explained more in details in Sec. 3, is energy dependent, therefore its magnitude changes depending on the energy of the interaction. The QCD

coupling constant $\alpha_s = \frac{g_s^2}{4\pi}$ is commonly used instead of g_s . At high energies (short distances) α_s is weak, this feature is called asymptotic freedom because in the asymptotic limit of infinite energy the particles behave as free particles, thus the strong interaction vanishes. On the other way round at low energy (large distances) α_s is very strong and the quarks are bounded together. An interesting characteristic of QCD is the color confinement: even at high energies quarks can not be isolated, they will combine in mesons or baryons to form a neutral color state and a quark carrying a color charge can not be observed. The physical explanation of color confinement is related to the self interaction of gluons, but there is no analytical explanation of this phenomenon.

2.2 $SU(2)_L \times U(1)_Y$

$SU(2)_L \times U(1)_Y$ is the symmetry group of the electroweak interactions, it has four generators, I^a with $a = 1, 2, 3$ for $SU(2)_L$ and Y for $U(1)_Y$, forming the algebra:

$$[I^a, I^b] = i\epsilon^{abc}I^c, [I^a, Y] = 0 \quad (2.6)$$

where ϵ^{abc} , the structure constants of $SU(2)$, are represented by the Levi-Civita anti-symmetric tensor. In the fundamental representation the generators $I^a = \frac{\sigma^a}{2}$ are the Pauli matrices.

Left-handed fermions appear as the fundamental representation of the group, right-handed fermion fields are invariant under I^a . The fermions having left chirality have weak isospin $\frac{1}{2}$ and form doublets while the right handed fermions have $I_3 = 0 \Rightarrow I = 0$ and are in singlets. For example for leptons:

$$\begin{pmatrix} \nu_l \\ l \end{pmatrix}_L, \quad l_R \quad l = e, \mu, \tau. \quad (2.7)$$

All the fermions carry the hypercharge Y coming from the Abelian group $U(1)_Y$. This theory contains two different couplings, g' associated to the gauge group $U(1)_Y$ and g related to the gauge group $SU(2)_L$. The gauge invariant Lagrangian of the electroweak gauge fields reads:

$$\mathcal{L}_{EW} = -\frac{1}{4}W_{\mu\nu}^a W_a^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \quad (2.8)$$

where

$$\begin{aligned} W_{\mu\nu}^a &= \partial_\mu W_\nu^a - \partial_\nu W_\mu^a + g\epsilon_{abc}W_\mu^b W_\nu^c, \\ B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu. \end{aligned} \quad (2.9)$$

W_μ^a appears as the adjoint representation of $SU(2)_L$ and is a triplet of vector fields. B_μ invariant under $U(1)_Y$ has the same structure as the photon field. The fermion-gauge

2 The Standard Model

field interaction is specified by

$$\begin{aligned}\mathcal{L}_F &= \sum_{\psi_L} \bar{\psi}_L i\gamma^\mu D_\mu^L \psi_L + \sum_{\psi_R} \bar{\psi}_R i\gamma^\mu D_\mu^R \psi_R, \\ D_\mu^{L,R} &= \partial_\mu - ig I_a^{L,R} W_\mu^a + ig' \frac{Y}{2} B_\mu, \\ I_a^L &= \frac{1}{2} \sigma_a, \quad I_a^R = 0\end{aligned}\tag{2.10}$$

where ψ_L and ψ_R , the left-handed and right-handed fermion fields, are a doublet and a singlet respectively. $I^{R,L}$ and Y are the weak isospin and hypercharge operators. The covariant derivative $D^{L,R}$ takes account the kinematic part of the fields and the interaction with the two vector fields.

The Lagrangian in eq. (2.8) unifies electromagnetic and weak interactions, it was proposed for the first time by Weinberg [20] and Salam [21]. Its power is to be gauge invariant in contrast to the one proposed by Glashow [22], due to the fact that it does not contain mass terms, which will arise after the spontaneous break of $SU(2)_L \times U(1)_Y$ through the Higgs mechanism. The fundamental fermions of the SM can be grouped as multiplets of $SU(3) \times SU(2)_L \times U(1)_Y$ as follows

$$\begin{aligned}& \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L, e_R, \mu_R, \tau_R \\ & \begin{pmatrix} u_r \\ u_g \\ u_b \end{pmatrix}, \begin{pmatrix} c_r \\ c_g \\ c_b \end{pmatrix}, \begin{pmatrix} t_r \\ t_g \\ t_b \end{pmatrix}, \begin{pmatrix} u_r \\ u_g \\ u_b \end{pmatrix}_R, \begin{pmatrix} d_r \\ d_g \\ d_b \end{pmatrix}_R, \begin{pmatrix} c_r \\ c_g \\ c_b \end{pmatrix}_R, \begin{pmatrix} s_r \\ s_g \\ s_b \end{pmatrix}_R, \begin{pmatrix} t_r \\ t_g \\ t_b \end{pmatrix}_R, \begin{pmatrix} b_r \\ b_g \\ b_b \end{pmatrix}_R\end{aligned}\tag{2.11}$$

In the SM particle content the neutrinos are not obtaining mass after the symmetry breaking and therefore they are massless. Moreover, only left-handed neutrinos are predicted by the theory.

2.3 Symmetry breaking mechanism

In this section we introduce the spontaneous breaking of $SU(2)_L \times U(1)_Y$, which leaves the $U(1)_{EM}$ (electromagnetic Abelian subgroup) unbroken and gives masses to the weak vector bosons Z^0 and W^\pm . Quarks and leptons get their masses via Yukawa interactions. This mechanism, often referred as the Higgs mechanism or electroweak symmetry breaking mechanism, was developed by Englert and Brout [23], Higgs [24] and Guralnik et al.[25] independently in the same year. The mechanism explains how the particles get their masses after the spontaneous breaking of a symmetry. Let us

2.3 Symmetry breaking mechanism

review briefly the mechanism. The Higgs Lagrangian reads:

$$\begin{aligned}\mathcal{L}_h &= (D_\mu \Phi)^\dagger (D^\mu \Phi) - V(\Phi), \\ D_\mu &= \partial_\mu - ig \frac{\sigma_a}{2} W_\mu^a + i \frac{g'}{2} B_\mu\end{aligned}\tag{2.12}$$

where:

$$\Phi(x) = \begin{pmatrix} \Phi^+(x) \\ \Phi^0(x) \end{pmatrix}\tag{2.13}$$

is a doublet of complex scalar fields with $Y = 1$ and four degrees of freedom. The covariant derivative D_μ takes into account the kinetic part of the field and its interaction with the gauge bosons, while $V(\Phi)$ is the potential developing a vacuum expectation value:

$$V(\Phi) = -\mu^2 \Phi^\dagger \Phi + \frac{\lambda}{4} (\Phi^\dagger \Phi)^2, \quad (\lambda, \mu^2 > 0)\tag{2.14}$$

where λ is the Higgs self coupling. The Lagrangian \mathcal{L}_h is symmetric under $SU(2)_L \times U(1)_R$ but the vacuum is not. The vacuum expectation value is not zero indeed:

$$\langle 0 | \Phi | 0 \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \quad v = \frac{2\mu}{\sqrt{\lambda}}\tag{2.15}$$

which leads to the spontaneous breaking of the symmetry. Nevertheless the vacuum is symmetric under transformation of $U(1)_{EM}$ preserving the electromagnetic symmetry. In polar coordinates the Higgs doublet reads

$$\Phi(x) = \frac{1}{\sqrt{2}} e^{i \frac{\chi_a(x)}{v} \sigma^a} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}\tag{2.16}$$

where $\chi_a(x)$, with $a = 1, 2, 3$, are three Goldstone boson fields. After the $SU(2)_L \times U(1)_Y$ symmetry breaking, the Goldstone bosons degrees of freedom mix with the W^\pm and Z^0 bosons and disappear, therefore the scalar doublet can be expressed as

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}\tag{2.17}$$

where $h(x)$ is the physical Higgs scalar field. The vector bosons W^\pm and Z and Higgs boson masses arise after the breaking and can be expressed as:

$$\begin{aligned}m_h &= \sqrt{2\lambda}v = \mu\sqrt{2}, \quad m_Z = \frac{\sqrt{g^2 + g'^2}}{2}, \\ m_W &= m_Z \cos(\theta_W) = \frac{gv}{2}, \quad \cos(\theta_W) = \frac{g}{\sqrt{g^2 + g'^2}}\end{aligned}\tag{2.18}$$

where $\cos(\theta_W)$ is the cosine of the electroweak mixing angle θ_W . Z^0 and W^\pm bosons were observed experimentally for the first time in 1983 [26–29] while for the Higgs boson

2 The Standard Model

the physics community had to wait until 2012 [1, 2]. The experimental measurements of their masses yields the values [3]

$$m_H = 125.10 \pm 0.14 \text{ GeV}, m_Z = 91.1876 \pm 0.0021 \text{ GeV}, m_Z = 80.379 \pm 0.012 \text{ GeV} \quad (2.19)$$

The Higgs potential in unitarity gauge can be rewritten as:

$$V(h) = \frac{m_h^2}{2}h^2 + \frac{m_h^2}{2v}h^3 + \frac{m_h^2}{8v^2}h^4. \quad (2.20)$$

The second and third terms on the r.h.s. of eq. (2.20) describe the Higgs triple and quartic self interactions. Quarks and leptons (with the exception of the neutrinos) get their masses after the symmetry breaking as well. Within the unitarity gauge the Yukawa Lagrangian stands:

$$\mathcal{L}_Y = - \sum_f m_f \bar{\psi}_f \psi_f - \sum_f \frac{m_f}{v} \bar{\psi}_f \psi_f h \quad (2.21)$$

where the index f runs over quarks and leptons. The first term on the r.h.s. is related to the fermion mass obtained after the symmetry breaking while the second term on the r.h.s. describes the interactions of the Higgs boson with fermions. The coupling of this interaction is $y_f = m_f/v$, considering the vacuum expectation value $v = 256 \text{ GeV}$ the only fermion whose interaction is not suppressed is the top quark. The top Yukawa coupling y_t measurement, normalised to SM value yields $1.26^{+0.31}_{-0.26}$ [30]. The measurement within the error is in agreement with the SM but the main value deviates from the SM prediction leaving room for New Physics extension.

2.4 The Higgs boson self couplings

In this Section we discuss the trilinear λ and quartic λ_4 Higgs boson self couplings. Part of the Section is adapted from [31] where the author of this thesis is a coauthor.

The trilinear Higgs boson coupling λ is of particular interest because it is directly connected to the Higgs potential. Deviations from the SM prediction would mean New Physics effects and give hints to extend the SM. The cross sections of the processes this coupling is involved in are indeed very small, for example the SM next-to-leading order (NLO) QCD cross section for Higgs boson pair production at 14 TeV is 32.95 fb. Therefore there is not a direct measurement of it but only constraints. The trilinear Higgs boson coupling λ can be constrained by measurements of Higgs boson pair production [32, 33], where the gluon fusion channel yields the largest cross section, and the ATLAS 95% confidence level (CL) limits on the total $gg \rightarrow HH$ cross section at $\sqrt{s} = 13 \text{ TeV}$ are currently $\sigma_{\text{max}}^{hh} = 6.9 \times \sigma_{SM}^{hh}$, where σ_{SM}^{hh} is the SM prediction, constraining trilinear coupling modifications to the range [32]

$$-5.0 \leq \lambda/\lambda_{SM} \leq 12.0 \quad (2.22)$$

where $\lambda_{\text{SM}} = \frac{m_h^2}{2v}$. The CMS 95% CL limits on the total cross section at $\sqrt{s} = 13 \text{ TeV}$ are looser than the ATLAS ones, $\sigma_{\text{max}}^{HH} = 22.2 \times \sigma_{\text{SM}}$, constraining trilinear coupling modifications to the range [33]

$$-11.8 \leq \lambda/\lambda_{\text{SM}} \leq 18.8. \quad (2.23)$$

The trilinear Higgs couplings can also be constrained in an indirect way, through measurements of processes which are sensitive to these couplings via electroweak loop corrections [34–45]. Such processes offer important complementary information, however they are susceptible to other BSM couplings entering the loop corrections at the same level, and therefore the limits on $c_{hhh} = \lambda/\lambda_{\text{SM}}$ extracted this way may be more model dependent than the ones extracted from the direct production of Higgs boson pairs. A corresponding experimental analysis based on single Higgs boson production processes has been performed [46], and recently combined constraints from single and double Higgs boson production became available [47]. Under the assumption that all deviations from the SM expectation are stemming from a modification of the trilinear coupling, the derived bounds on c_{hhh} at 95% CL from the combined analysis are $-2.3 \leq c_{hhh} \leq 10.3$ [47]. However once the couplings to vector bosons and/or fermions are allowed to vary as well, these bounds are diluted.

The idea of indirect constraints through loop corrections also has been employed trying to constrain the quartic Higgs boson self-coupling from (partial) EW corrections to Higgs boson pair production [48, 49].

Theoretical constraints on c_{hhh} are rather loose if derived in a largely model independent way. Recent work based on general concepts like vacuum stability and perturbative unitarity suggests that $|c_{hhh}| \lesssim 4$ for a new physics scale in the few TeV range [50–53]. More specific models can lead to more stringent bounds, see e.g. Refs. [54–58]. Recent phenomenological studies about the precision that could be reached for the trilinear coupling at the (HL-)LHC and future hadron colliders are summarised in Refs. [59–61].

The Higgs quartic coupling λ_4 , the other leading term of the Higgs potential, in the SM formulation is connected to the trilinear one by the relation $\lambda_{\text{SM}}/\lambda_{4,\text{SM}} = v$. This coupling appears at LO in the production of three Higgs bosons. The leading order cross section of the process, for $m_H = 125 \text{ GeV}$ is 0.040 fb at 14 TeV [62], too small to be observed at current particles colliders. Therefore there are no constraints on the quartic Higgs coupling. The Higgs potential and the Higgs mass are related to vacuum stability. The running of the quartic coupling gives information about the nature of the electroweak vacuum. If λ_4 becomes negative at scales much below the Plank scale $M_P = 1.220910 \times 10^{19} \text{ GeV}$, the Higgs field could tunnel from the actual metastable vacuum to a true vacuum. By the way, the beta function of the Higgs quartic coupling $\beta_{\lambda_4}(\lambda_4, y_t, g_s)$ is very sensitive of variation of its arguments, and the actual measurements of the Higgs boson mass and in particular the top quark mass, do not allow to answer the question whether the SM vacuum is stable or not [63].

3 Higher order perturbation theory

In this Chapter we review the fundamental concepts to carry out the theoretical computations used to describe phenomena at particles colliders, focusing on hadron collisions. The content of this section is adapted from [64–68].

3.1 Factorization theorem

The Large Hadron Collider (LHC) is a proton-proton collider. The energies are so high that the interactions are no more between the two protons, but among partons i.e. gluons and quarks. The factorization theorem [69] allows to write the cross section for the scattering of two hadrons h_1, h_2 producing the final state X in the following way

$$\sigma_{h_1 h_2 \rightarrow X} = \sum_{i,j} \int_0^1 dx_1 \int_0^1 dx_2 f_{h_1,i}(x_1, \mu_F) f_{h_2,j}(x_2, \mu_F) \cdot \hat{\sigma}_{ij \rightarrow X}(\hat{s}, \alpha_S(\mu_R), \mu_F, \mu_R). \quad (3.1)$$

h_1 and h_2 are the two interacting hadrons, i, j are the partons indices and the integration variables x_1, x_2 represent the hadrons' momenta fractions of the two partons. $f_{h_1,i}(x_1, \mu_F), f_{h_2,j}(x_2, \mu_F)$ are the partonic distribution functions (PDFs) and are related to the probability to find two partons i, j with momenta fractions x_1, x_2 at the energy μ_F inside the protons. PDFs can not be computed theoretically, they are determined by fits of experimental data. They are the universal part of the formula in the sense that do not depend on the specific process. Even though PDFs are non perturbative objects, their evolution with scale μ_F can be computed perturbatively according to the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations [70–72]. In fig. 3.1 the comparison among PDFs fits is shown, for gluons on the left and up quarks on the right. In the plots are represented three sets of PDFs CT14 [73], MMHT [74] and NNPDF [75], normalised to the central value of the CT14 distribution, as functions of x , the momentum fraction carried out by the partons, at next-to-next to leading order (NNLO) in the DGLAP expansion. The energy scale is $Q^2 = 100 \text{ GeV}$. These three sets are combined by the PDF4LHC Working Group, providing the PDF4LHC15 set [76] which is suitable for LHC simulations.

$\hat{\sigma}_{ij}$ is the partonic cross section for the partons i, j , the process dependent part of (3.1), with $\sqrt{\hat{s}} = \sqrt{x_1 x_2 s}$ partonic centre of mass energy. Eq. (3.1) can be interpreted as the sum over all the possible partons, integrated over all the possible momentum fractions of the partonic cross sections weighted by the PDFs. Once the partonic cross sections are known eq. (3.1) allows to calculate the total cross section for the final state X .

3 Higher order perturbation theory

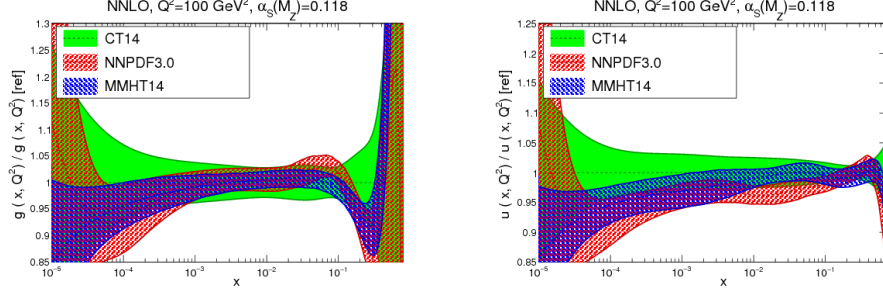


Figure 3.1: Comparison of the gluon and up quark PDFs at energy scale $Q^2 = 100 \text{ GeV}$, normalized to the central value of CT14 [76].

3.2 QCD perturbative expansion

The partonic cross section is calculated by the perturbative expansion in the strong coupling constant,

$$\hat{\sigma} = \alpha_S^n(\hat{\sigma}^{(0)} + \alpha_S \hat{\sigma}^{(1)} + \alpha_S^2 \hat{\sigma}^{(2)} + \dots) \quad (3.2)$$

where the first term in brackets represents the leading order (LO), the second the next-to-leading order (NLO), the third the next-to-next to leading order (NNLO) and so on. n is the strong coupling power of the LO cross section. We are allowed to take the expansion in eq. (3.2) when $\alpha_S \ll 1$, which ensures that higher order terms of the expansion can be neglected when sufficiently small. Λ_{QCD} is the QCD scale parameter, defining the value of energy scale at which the coupling constant diverges. The QCD perturbative expansion can be performed when computing $\hat{\sigma}$ at energies much larger than $\Lambda_{QCD} \sim 200 \text{ MeV}$ roughly the pion mass. NLO and NNLO corrections are often very important to reach the theoretical precision needed to match the experimental results.

Fig. 3.2 shows the cross section for $t\bar{t}$ production as function of the partonic invariant mass $\sqrt{\hat{s}}$, highlighting the importance of NNLO corrections. The LO and NLO error bars are too large and do not allow to make reliable predictions, while the NNLO one is thinner with magnitude of the uncertainty of the same order of the experimental data. In general the cross section calculation has to be performed at orders higher than LO, in order to provide precise predictions.

We can write the partonic cross section as

$$\hat{\sigma} = \int_{\Phi} d\Phi |\mathcal{M}|^2 \quad (3.3)$$

thus the integral over the phase space $d\Phi$ of the scattering amplitude squared and study the perturbative expansion of $|\mathcal{M}|^2$ itself:

$$|\mathcal{M}|^2 = \alpha_S^n (|\mathcal{M}_0|^2 + \alpha_S |\mathcal{M}_1|^2 + \alpha_S^2 |\mathcal{M}_2|^2 + \dots). \quad (3.4)$$

3.3 Differential cross section distributions

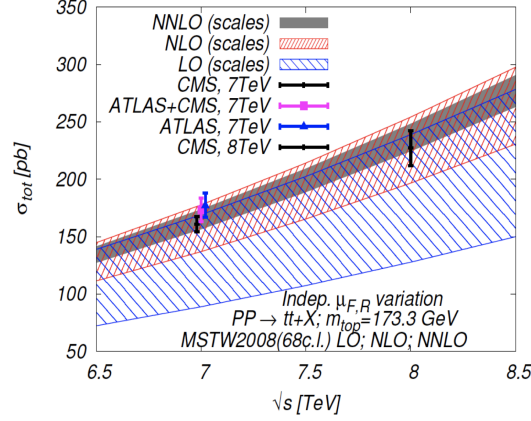


Figure 3.2: $t\bar{t}$ total cross section as function of the invariant mass. The picture shows the experimental data (black vertical lines) and the theoretical predictions (color bars) at different orders [77].

3.3 Differential cross section distributions

An important observable in the study of a process is the differential cross section distribution, defined as the derivative of the cross section with respect to a physical quantity. For example considering a two-to-two partonic process $i_1 i_2 \rightarrow f_1 f_2$, where i_1 and i_2 represent the initial state particles and f_1 and f_2 the final state particles, an interesting distribution for phenomenological studies is

$$\frac{d\hat{\sigma}}{dm_{f_1 f_2}}, \quad (3.5)$$

where $m_{f_1 f_2}^2 = (p_{f_1} + p_{f_2})^2 = \hat{s}$ is the invariant mass squared of the final state. Recalling the momentum conservation $(p_{f_1} + p_{f_2})^2 = (p_{i_1} + p_{i_2})^2$, where p_k the momentum of the particle k , $m_{f_1 f_2}^2$ is equal to the invariant mass squared of the initial state $m_{i_1 i_2}^2$. The $m_{f_1 f_2}$ differential distribution allows to study the partonic cross section as function of the invariant mass and roughly speaking it gives the value of the partonic cross section at the energy $m_{f_1 f_2}$. Another important differential distribution is the transverse momentum distribution

$$\frac{d\hat{\sigma}}{dp_{T,f_1}}. \quad (3.6)$$

where $p_{T,f_1} = \sqrt{p_{f_1,x}^2 + p_{f_1,y}^2}$ is the transverse momentum of the particle f_1 , i.e. the component of the three-momentum perpendicular to the beam axis. Differential distributions can be studied at different orders of the perturbative expansion analogously to the total cross section.

Considering hadronic initial states, the differential cross section distribution in the quantity Y can be obtained from the partonic ones using the factorization theorem

3 Higher order perturbation theory

similarly to eq. (3.1)

$$\frac{d\sigma_{h_1 h_2 \rightarrow X}}{dY} = \sum_{i,j} \int_0^1 dx_1 \int_0^1 dx_2 f_{h_1,i}(x_1, \mu_F) f_{h_2,j}(x_2, \mu_F) \cdot \frac{d\hat{\sigma}_{ij \rightarrow X}(\hat{s}, \alpha_S(\mu_R), \mu_F \mu_R)}{dY}. \quad (3.7)$$

3.4 Regularization and renormalisation

The LO scattering amplitudes are usually given by Feynman tree diagrams, but as discussed in the previous section the LO calculations are not enough to provide reliable predictions and higher accuracy is needed. This leads to the computation of higher order diagrams. These involve loop integrals which, as we are going to show, can present two kinds of divergencies. Let us consider the simplest loop integral, the so called bubble integral, shown in Fig. 3.3

$$B_0 = \int \frac{d^4 k}{(2\pi)^4} \frac{1}{[k^2 - m^2 + i\delta][(k+p)^2 - m^2 + i\delta]} \quad (3.8)$$

where p is the momentum entering the bubble, m the mass of the two particles in the loop and k the loop momentum, the variable we are integrating over.

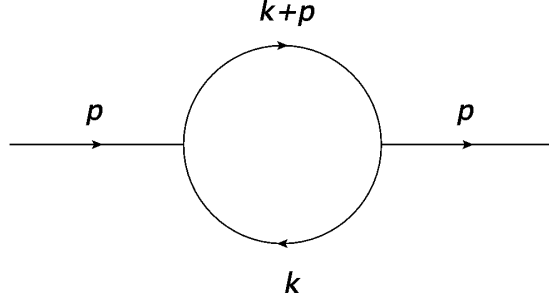


Figure 3.3: The bubble integral diagram, where p is the momentum entering and exiting the loop and k the loop momentum.

At the upper integration limit, the integral behaves as

$$B_0 \sim \int^\infty \frac{d|k| |k|^3}{|k|^4} = \int^\infty \frac{d|k|}{|k|} \quad (3.9)$$

which means the integral diverges logarithmically. $|k|$ is the euclidian module of the loop momentum. The divergence in the high energy limit is called ultraviolet (UV) divergence. At the lower integration limit $|k| \rightarrow 0$ when $m = 0$ a divergence occurs as well, in this case we are talking of infrared (IR) divergences.

In general loop integrals have these two kinds of divergencies which need to be fixed in order to give physical meaning to the high order terms of the perturbative expansion.

The loop-integral divergencies are treated by the dimensional regularization, where the integral is regularized in D dimension and the divergences become poles in $\epsilon = (4 - D)/2$. In the case of UV divergencies the poles are absorbed in the redefinition of the Lagrangian, while one can see that the IR poles cancel at cross section level when they are averaged over all the degenerate initial and final states, except for initial state collinear divergencies which are absorbed in the PDFs redefinition.

3.4.1 Dimensional regularization

Dimensional regularization (DR) [78] consists in an analytical continuation in the space-time dimension. The integrals are not evaluated in 4 dimensions but in $D = 4 - 2\epsilon$ dimensions. For $\epsilon > 0$ the loop integrals do not have UV divergencies and the integration can be carried out. The dimensional regularization keeps the integrals gauge and Lorentz invariant and so does not modify properties as the Ward identities. Even more this approach is the only one which allowed to prove the renormalisability of non-abelian gauge theories as QCD maintaining the gauge symmetry. The DR implies the coupling rescaling $g \rightarrow g\mu_R^\epsilon$, to ensure that all the terms in the Lagrangian have the right mass dimension. Working in D dimensions, $g_{\mu\nu}g^{\mu\nu} = D$, and the Clifford algebra of the Dirac matrices has to be modified consistently:

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu} \quad \text{with } g_{\mu\nu}g^{\mu\nu} = D. \quad (3.10)$$

The matrix $\gamma_5 = i\gamma_0\gamma_1\gamma_2\gamma_3$ is not obvious and uniquely defined to continue in D dimensions. The way to treat γ_5 and the helicities of external and internal particle fields leads to the choice of a different regularization scheme. The scheme used in the calculations of Chapters 6 and 8 is "conventional dimensional regularization" (CDR), where both the internal and external fields are treated as D -dimensional.

Coming back to the bubble example, the integral after dimensional regularization can be solved as

$$B_0 = \int \frac{d^D k}{(2\pi)^D} \frac{1}{[k^2 - m^2 + i\delta][(k+p)^2 - m^2 + i\delta]} = \frac{1}{\epsilon} B_{div} + B_{finite} + \mathcal{O}(\epsilon) \quad (3.11)$$

where the pole term $\frac{1}{\epsilon} B_{div}$ is the piece containing the divergence and B_{finite} the finite part of the integral. The analytical expression of the integral has been calculated in [79].

3.4.2 Renormalisation

Here we describe the procedure to absorb the UV poles redefining the Lagrangian parameters, the so-called renormalization. Let us define three Lagrangians:

$$\mathcal{L}_0(A_0, q_0, m_0, g_0), \quad \mathcal{L}_P(A, q, m, g\mu_R^\epsilon), \quad \mathcal{L}_C(A, q, m, g\mu_R^\epsilon) \quad (3.12)$$

where $\mathcal{L}_0, \mathcal{L}_P$ and \mathcal{L}_C are the bare, physical and counterterm Lagrangians. All of them specify a general theory with coupling g , vector boson fields A and fermion fields q of

3 Higher order perturbation theory

mass m . After defining the counter terms relating the bare and physical quantities via

$$g_0 = \mu_R^\epsilon Z_g g = g + \delta g, \quad m_0 = Z_m m = m + \delta m, \quad q_0 = Z_q^{\frac{1}{2}} q, \quad A_0^\mu = Z_A^{\frac{1}{2}} A \quad (3.13)$$

where $Z_X = 1 + \delta Z_X$ are the renormalisation constants, the bare Lagrangian can be written as a combination of the other two

$$\mathcal{L}_0(A_0, q_0, m_0, g_0) = \mathcal{L}_P(A, q, m, g\mu_R^\epsilon) + \mathcal{L}_C(A, q, m, g\mu_R^\epsilon). \quad (3.14)$$

The coupling g is now dependent on the renormalization scale μ_R , and it can be shown that this yields the coupling to be energy dependent [80, 81]; its running with the energy is described by the β function.

The renormalization concept is related to the fact that the parameters of the bare Lagrangian are not the one described by the fields, they are not observable and thus unphysical. The relation between bare and physical parameters is given by the counterterms. In other words only combining the bare and counterterms Lagrangians one can obtain the physical and observable Lagrangian \mathcal{L}_P and in the combination of the two the UV poles are canceled.

3.4.3 Infrared divergencies

The KLN(Kinoshita-Lee-Nauenberg) theorem [82, 83] states that at cross section level the IR divergencies of a Feynman diagram cancel when summing it with all its degenerate diagrams, the so-called real radiation. At NLO degenerate diagrams are diagrams with an extra state yielding a cross section indistinguishable from the LO cross section in experiments. The extra state can be a massless particle with momentum close to zero (soft divergencies) or collinear with the momentum of one of the initial/final state particles (collinear divergencies). The cross section becomes finite, i.e. the IR divergences cancel, only when summing over all the possible final state multiplicities. Considering the UV divergencies renormalized, the NLO cross section of a given process can be written as

$$\sigma_{NLO} = \int_{n+1} d\sigma_R + \int_n d\sigma_V \quad (3.15)$$

where the first term on the r.h.s, integrated over the $n+1$ particles phase space, is the piece related to the degenerate diagrams indistinguishable from LO kinematics, which yields divergencies after the integration. The second term on the r.h.s., integrated over the n particles phase space, is the virtual correction presenting the IR divergencies and it needs to be regularized to carry out the integration. One can see that, if the procedure is performed correctly, the poles cancel each other with the exception of remaining initial state collinear singularities. They are absorbed by renormalizing the bare PDFs presented in eq. (3.1) in the redefinition $f_{h_k,i}^0(x_1) \rightarrow f_{h_k,i}(x_1, \mu_F)$. Similarly to the renormalization of g , where the coupling develops a dependence on the renormalization scale μ_R , the PDFs develop a dependence on the factorization scale μ_F .

3.4 Regularization and renormalisation

A good technique used to deal with the IR divergencies at NLO is the Catani-Seymour method [84]. The first step is to add and subtract a subtraction term:

$$\sigma_{NLO} = \int_{n+1} [d\sigma_R - d\sigma_A] + \int_n \left[d\sigma_V + \int_1 d\sigma_A \right] \quad (3.16)$$

where $d\sigma_A$ approximates in D dimensions the singular behavior of $d\sigma_R$. In the first term on the r.h.s., one can perform the limit $\epsilon \rightarrow 0$ in the integrand removing the poles of the quantity $[d\sigma_R - d\sigma_A]$ which now is finite and the integration can be carried out in 4 dimensions. In the second term on the r.h.s., the integration of $d\sigma_A$ over the one-parton subspace yields the IR poles in ϵ which cancel the ones of $d\sigma_V$, then the limit $\epsilon \rightarrow 0$ can be taken and the integration over the n partons phase space can be carried out. The general form of the subtraction term is given by

$$d\sigma_A = \sum_{dipoles} d\sigma_B \otimes dV_{dipole} \quad (3.17)$$

where dV_{dipole} representing the dipole factors are universal and independent on the process, while $d\sigma_B$ is the the color and spin projection of the Born-level exclusive cross section, so process dependent. \otimes stands for properly defined phase space convolutions and sums over colour and spin indices. We can then rewrite eq. (3.16) as

$$\sigma_{NLO} = \int_{n+1} \left[(d\sigma_R)_{\epsilon=0} - \left(\sum_{dipoles} d\sigma_B \otimes dV_{dipole} \right)_{\epsilon=0} \right] + \int_n [d\sigma_V + d\sigma_B \otimes \mathbf{I}]_{\epsilon=0} \quad (3.18)$$

where $\mathbf{I} = \sum_{dipoles} \int_1 dV_{dipole}$, the summed integration of the dipole factors over the one-parton phase space, contains the ϵ poles canceling those of $d\sigma_V$. The NLO Catani Seymours subtraction terms can be computed following the procedure described in [84]. This section conclude the Chapter where we reviewed the main concepts regarding the high energy calculations, focusing on QCD processes.

4 Effective Field Theories formalism

An Effective Field Theory (EFT) is an approximate theory used to parametrise phenomena at a chosen energy scale, ignoring the fundamental structure and the degrees of freedom at higher energies. The predictions of such kind of theory are more accurate far from the underlying dynamics scale Λ . One example is the Higgs Effective Field Theory (HEFT) where the energy scale $\Lambda = m_t$ is the top quark mass. The HEFT approximation, where the $m_t \rightarrow \infty$ limit is taken, integrates out the top quark dependence in the loop integrals. It provides a good description of the SM and Higgs Physics for energies below m_t . An example of HEFT diagram for Higgs boson pair production is shown in Fig. 4.1, where the top loop is shrunk resulting in an effective gluon-Higgs interaction specified at LO by the Lagrangian :

$$\mathcal{L}_{effective} \supset \frac{\alpha_S}{12\pi} \frac{h^2}{v^2} G_{\mu\nu}^a G_a^{\mu\nu}. \quad (4.1)$$

Several models aiming to extend the physics beyond the SM, parametrise the New Physics using an EFT where Λ is related to the masses of new particles.

Such a framework can be formulated in various ways, where we can distinguish two main categories, often called “linear EFT” and “non-linear EFT”. The linear EFTs [85, 86], also known as “SMEFT” [87–91], are organised by canonical dimensions, formulated as power series in the dimensionful parameter $1/\Lambda$. The non-linear EFTs are organised by chiral dimensions. The corresponding formalism, including a light Higgs boson, has been developed in Refs. [92–105] and usually goes by the name “Electroweak Chiral Lagrangian” (EWChL). The calculations presented in Chapters 6,8 and 7 are performed within the “non-linear EFT” framework, where the Higgs field is an electroweak singlet. The main benefit of this approach is that the anomalous

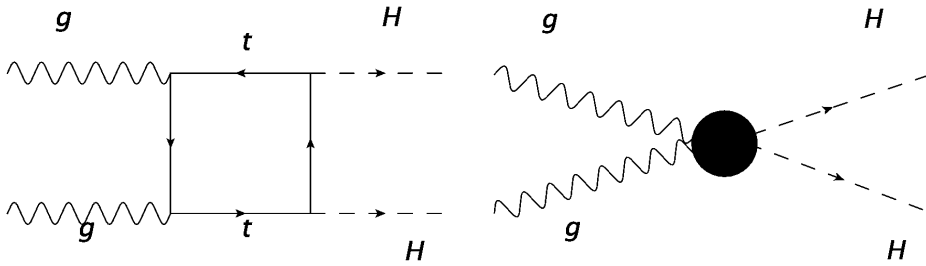


Figure 4.1: Higgs boson pair production LO diagrams. On the left the diagram of the full theory, on the right the same diagram in the HEFT approximation where the top loop is shrunk.

Higgs couplings are singled out, in a systematic way, as the dominant New Physics effects in the electroweak sector. Sometimes the electroweak chiral Lagrangian with a light Higgs boson is also referred to as *Higgs Effective Field Theory (HEFT)* in the literature. The two EFTs are unrelated and should be carefully distinguished. Here we employ the term *electroweak chiral Lagrangian* for the non-linear EFT of physics beyond the SM, and reserve the expression *HEFT* for the heavy-top limit in Higgs interactions. In this chapter we describe the two possible realisations of an EFT. We focus on the non-linear realisation but first we give a general overview on the linear one.

4.1 Linear EFT

The linear EFT framework parametrising New Physics is the so called SM Effective Field Theory (SMEFT), where to the SM Lagrangian are added new operators with canonical dimension larger than 4. The new Lagrangian has the same $SU(3)_C \times SU(2)_L \times U(1)_Y$ symmetry of the SM. Generally the SMEFT Lagrangian reads [106]:

$$\mathcal{L} = \mathcal{L}_{SM} + \sum_i \frac{c_i^{(5)}}{\Lambda} \mathcal{O}_i^{(5)} + \sum_i \frac{c_i^{(6)}}{\Lambda^2} \mathcal{O}_i^{(6)} + \sum_i \frac{c_i^{(7)}}{\Lambda^3} \mathcal{O}_i^{(7)} + \sum_i \frac{c_i^{(8)}}{\Lambda^4} \mathcal{O}_i^{(8)} + \dots, \quad (4.2)$$

where $\mathcal{O}_i^{(D)}$ is the $SU(3)_C \times SU(2)_L \times U(1)_Y$ operator of canonical dimension D and $c_i^{(D)}$ the related Wilson coefficient containing the information about New Physics. SMEFT aims to parametrise BSM theory effects at energies much lower than Λ , the scale where the BSM particles can be produced. Given v , the vacuum expectation value, the contribution of each $\mathcal{O}_i^{(D)}$ scales as $(v/\Lambda)^{(D-4)}$ when the energy scale is about the order of v . By construction $v/\Lambda \ll 1$ therefore SMEFT corrections can only yield small deviations from SM predictions.

All D=5 operators violate lepton number [107] and in general all odd-D operators violate baryon number minus lepton number (B-L) [108] and experimental constraints oblige their Wilson coefficients to be suppressed at a level which makes them unobservable at the LHC [109]. For this reason they usually are not taken into account. Moreover considering that D=8 operators are suppressed by $(v/\Lambda)^4$ in several works only the dominant dimension 6 operators are considered.

4.2 Non-linear EFT

In this Section we review the non-linearly realised Effective Field Theory, the EWChL including a light Higgs boson [101, 103, 110]. The Section is adapted from Ref. [111] where the author of this thesis is a coauthor.

This framework provides us with a consistent EFT for New Physics in the Higgs sector, as we will summarize in the following.

To leading order the Lagrangian is given by

$$\begin{aligned}
\mathcal{L}_2 = & -\frac{1}{2}\langle G_{\mu\nu}G^{\mu\nu}\rangle - \frac{1}{2}\langle W_{\mu\nu}W^{\mu\nu}\rangle - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} + \sum_{\psi=q_L, l_L, u_R, d_R, e_R} \bar{\psi}i\not{D}\psi \\
& + \frac{v^2}{4} \langle D_\mu U^\dagger D^\mu U \rangle (1 + F_U(h)) + \frac{1}{2}\partial_\mu h \partial^\mu h - V(h) \\
& - v \left[\bar{q}_L \left(Y_u + \sum_{n=1}^{\infty} Y_u^{(n)} \left(\frac{h}{v} \right)^n \right) U P_+ q_R + \bar{q}_L \left(Y_d + \sum_{n=1}^{\infty} Y_d^{(n)} \left(\frac{h}{v} \right)^n \right) U P_- q_R \right. \\
& \left. + \bar{l}_L \left(Y_e + \sum_{n=1}^{\infty} Y_e^{(n)} \left(\frac{h}{v} \right)^n \right) U P_- l_R + \text{h.c.} \right]. \tag{4.3}
\end{aligned}$$

The first line is the unbroken SM written with different notation with respect of Chap. 2, the remainder represents the Higgs sector. Here h is the Higgs field and $U = \exp(2i\varphi^a I^a/v)$ encodes the electroweak Goldstone fields φ^a , with I^a the generators of $SU(2)$. v is the electroweak vacuum expectation value, $P_\pm = 1/2 \pm I_3$, and

$$D_\mu U = \partial_\mu U + igW_\mu U - ig'B_\mu U I_3. \tag{4.4}$$

The trace of a matrix A is denoted by $\langle A \rangle$. The left-handed doublets of quarks and leptons are written as q_L and l_L , the right-handed singlets as u_R , d_R , e_R . Generation indices are omitted. In the Yukawa terms the right-handed quark and lepton fields are collected into $q_R = (u_R, d_R)^T$ and $l_R = (0, e_R)^T$, respectively. In general, different flavour couplings $Y_{u,d,e}^{(n)}$ can arise at every order in the Higgs field h^n , in addition to the usual Yukawa matrices $Y_{u,d,e}$. The h -dependent functions are

$$F_U(h) = \sum_{n=1}^{\infty} f_{U,n} \left(\frac{h}{v} \right)^n, \quad V(h) = v^4 \sum_{n=2}^{\infty} f_{V,n} \left(\frac{h}{v} \right)^n. \tag{4.5}$$

In the limit where

$$f_{U,1} = 2, \quad f_{U,2} = 1, \quad f_{V,2} = f_{V,3} = \frac{m_h^2}{2v^2}, \quad f_{V,4} = \frac{m_h^2}{8v^2}, \quad Y_f^{(1)} = Y_f, \tag{4.6}$$

and all other couplings $f_{U,n}$, $f_{V,n}$, $Y_f^{(n)}$ equal to zero, the Lagrangian in (4.3) reduces to the usual SM. For generic values of those parameters, the Lagrangian describes the SM with arbitrary modifications in the Higgs couplings. While the deviations of these couplings from their SM values could, in principle, be of order unity, the parametrisation in (4.3) remains relevant as long as the anomalous Higgs couplings are the dominant New Physics effects at electroweak energies. Employing (4.3), makes the assumption that this is the case. A useful property of the Lagrangian (4.3) is therefore that it allows to concentrate on anomalous Higgs couplings in a systematic way [104, 112].

In fact, the intuitive picture of introducing (4.3) as the SM with modified Higgs couplings can be formulated as a consistent EFT. Because of the need to write the modified Higgs couplings in a gauge-invariant way, the Higgs field has to be represented

as an electroweak singlet h , independent of the Goldstone matrix $U = \exp(2i\varphi^a I^a/v)$. The latter transforms as $U \rightarrow g_L U g_Y^\dagger$ under the SM gauge group. The symmetry is non-linearly realised on the Goldstone fields φ^a . The Lagrangian (4.3) is then non-renormalisable (in the traditional sense) as it contains interaction terms of arbitrary canonical dimension. The EFT is therefore not organised by the canonical dimension of operators, but rather by *chiral counting* in analogy to the chiral perturbation of pions in QCD. Chiral counting is equivalent to an expansion in loop orders L , which can be conveniently counted by assigning *chiral dimensions* $d_\chi \equiv 2L + 2$ to fields and weak couplings. This assignment is simply 0 for bosons, and 1 for each derivative, fermion bilinear and weak coupling:

$$d_\chi(A_\mu, \varphi, h) = 0, \quad d_\chi(\partial, \bar{\psi}\psi, g, y) = 1. \quad (4.7)$$

Here A_μ represents a generic gauge field, φ the Goldstone bosons, and h the Higgs scalar. g denotes any of the SM gauge couplings g, g', g_s , and y any other weak coupling, such as the Yukawa couplings or the square-roots of the parameters $f_{V,n}$ in the Higgs potential.

Based on this counting, the leading-order expression (4.3) can be constructed from the SM field content and symmetries as the most general Lagrangian of chiral dimension 2. Leading processes are described by tree-level amplitudes from (4.3). Next-to-leading order effects come from one-loop contributions of (4.3) and from tree-level terms of the NLO Lagrangian \mathcal{L}_4 . Both are considered to be of ‘one-loop order’, or chiral dimension $d_\chi = 4$.

5 Machine Learning tools

Machine learning (ML) is a subfield of artificial intelligence (AI) where the algorithms aim to learn automatically from the data and to fit the data into models able to make predictions. In traditional computer science approaches the algorithms yielding the predictions are explicitly programmed, the coder writes the instructions and rules the algorithm has to follow to fit the data.

In ML the algorithms make the machine train on data sets and through statistical analysis find the best model fitting the data. These features of machine learning allow to build models in an easier way than using traditional computer science and are very powerful tools when the data are described by complex or unknown distributions. The machine learning algorithms are used for two tasks: regression (interpolation of an unknown function) and classification (collecting the outputs in two or more classes). Classification algorithms can be divided in two main types according to the learning method: supervised and unsupervised learning. In supervised learning the algorithms are trained on datasets labelled by humans, while in unsupervised learning the algorithms are fed by unlabelled data.

There are several algorithms available on the machine learning market such as Decision trees, K-nearest-neighbors, Support Vector Machine, K-Means, Principal Component Analysis, RandomForest, GradientBoosting and Neural Networks (NN). ML techniques are widely used by companies providing services for the every day life. Examples of ML based applications are: facial and voice recognition, virtual assistants, traffic predictions, social media services, search engines, product recommendations and so on.

Problems related to the analysis of big datasets are optimal to be approached by machine learning techniques and because of this ML plays an important role in Physics as well. In the last years ML applications to physics have become trending and the number of related papers has grown considerably. The machine learning techniques in high energy physics have been applied to constrain the EFT/new physics parameter space, [113–116], in jet and top quark identification [117–126], new physics searches [115, 116, 127–136] and PDFs [137]. In this Chapter we are presenting in more detail the machine learning tools we used in the calculation of Chapter 7.

5.1 Perceptron and Fully Connected Network

Neural Networks are a set of algorithms modeled after the human brain, designed to recognize patterns. The easiest and first application of a NN is the perceptron, a network where input and output are directly connected, and the output layer consists of only one neuron (or node). The perceptron is built up for a classification task. The network is shown in Fig. 5.1.

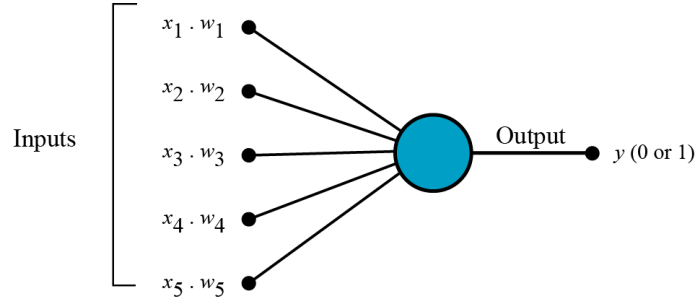


Figure 5.1: Perceptron architecture [138].

The input array \underline{x} is multiplied by the weights array \underline{w} , then a constant, the bias b , is added to the scalar product and passed through the activation function f in order to break the linearity. The output is $y = f(\underline{w} \cdot \underline{x} + b)$. The result of this operations is compared to a defined threshold value. If the activation function output is below threshold the input is of class "0" otherwise it is of class "1". If we consider the same model but without a threshold, we obtain a network where the output assumes continuous values, therefore a model for regression.

The most common used activation functions are:

- Linear function: $f(x) = x \rightarrow f(x) \in (-\infty, \infty)$
- Sigmoid function: $f(x) = 1/(1 + e^{-x}) \rightarrow f(x) \in [0, 1]$
- Hyperbolic tangent function: $f(x) = (e^x - e^{-x})/(e^x + e^{-x}) \rightarrow f(x) \in [-1, 1]$
- Rectified Linear Unit (ReLU) function : $f(x) = \max(0, x) \rightarrow f(x) \in [0, \infty)$.

Sigmoid and hyperbolic tangent functions are mainly used for binary classification. The linear function is barely used because it is not helpful when the complexity of the problem increases. The ReLU is the most used activation function because of its simplicity and it is very easy to compute its derivative. The activation function has to be differentiable because its gradient is used to update the weights via the gradient descent procedure.

The loss function (or cost function) is used to estimate how correctly the model predicts the inputs. It measures the distance between the predicted and expected outputs. The loss is then used to update the weights in order to improve the model performance. There are several candidates to pick as loss function, the choice depends on the problem and the feature we want to focus on. Usually Mean Squared Error (MSE) and Mean Absolute Error (MAE) functions are chosen for regression while Cross-Entropy is the most used for classification problems.

To obtain a model describing correctly the data the goal is to minimize the loss function reaching its global minimum. This results is achieved updating the weights

according to:

$$w'_i = w_i - \alpha \frac{\partial}{\partial w_i} Loss \quad (5.1)$$

where w_i is the i^{th} component of the weight array \mathbf{w} . α is the learning rate, an hyperparameter the coder has to tune. A value too large can lead to miss the global minimum of the loss function, while a value too small can over increase the computational time. At time zero the weights are randomly initialized, then when the training starts this procedure is iteratively repeated until the asked precision is reached.

In order to avoid the overfitting of the model the loss function can be regularized. There are 3 kinds of regularizations:

- L1 or Lasso Regression: $Loss \rightarrow Loss + \lambda_1 ||\mathbf{w}||$
- L2 or Ridge Regression: $Loss \rightarrow Loss + \lambda_2 ||\mathbf{w}||^2$
- Elastic Regression: $Loss \rightarrow Loss + \lambda_1 ||\mathbf{w}|| + \lambda_2 ||\mathbf{w}||^2$

where λ_1 and λ_2 are the hyperparameters used to regularize the weights. Within all the three kinds of regressions the weights are minimized along with the loss function. Regularization assumes that smaller weights generate simpler models and therefore it decreases the probability of overfitting.

The perceptron model can be extended using more neurons and adding more layers between the input and output, called hidden layers. This generalization is called Fully Connected Neural Network (FCNN) where all the neurons of a layer are connected to the neurons of the following layer. In Fig 5.2 is shown an example of FCNN with two layers. The input has 5 nodes (the dimension of the input array), the hidden layer has 4 nodes and the output 1 node. Each output a_j of the hidden layer's nodes is a linear transformation of the input array composed with the activation function. The procedure is reiterated and the a_j are linearly transformed and composed with the activation function to yield the output. The activation function, f and g in the example, can differ for the two layers.

This example shows that the working flow of the FCNN is the same of the perceptron, repeated layer by layer. Because of this, the back-propagation of the error has to be taken into account in order to update each weight according to its contribute to the error. Other types of Neural Networks are: Convolutional Neural Network (CNN), Recurrent Neural Network (RNN) and Recursive Neural Network (RNN). In all these networks the connections between the neurons are realized in more complex ways. When the hidden layers are more than one we talk of deep learning, otherwise the network is referred to as shallow network.

5.2 Autoencoders

NN models can be combined to achieve more sophisticated results, one particular case is the autoencoder. The autoencoder goal is to replicate the inputs, compressing them

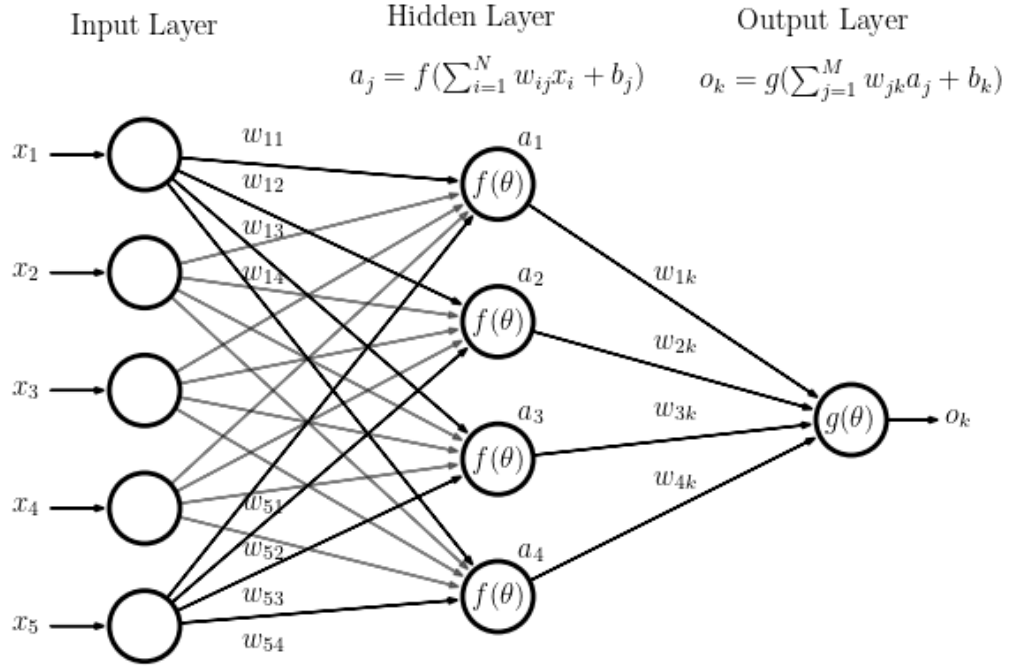


Figure 5.2: Architecture of a Fully Connected Network [139, 140].

in a latent-space representation and then reconstructing the outputs from this representation. The autoencoder architecture is shown in Fig. 5.3.

The network is divided in two parts: the encoder, compressing the data and the decoder decompressing it. The power of the autoencoder is not to copy the inputs into outputs, pretty useless task, but to obtain the latent representation. If the latent representation has smaller dimension of the input the autoencoder is called undercomplete, while on the other way round is called overcomplete. In the undercomplete configuration the algorithm is forced to learn the most important features of the data distribution. The autoencoder undercomplete, along with Principal Component Analysis, is the best tool for dimensionality reduction, allowing to isolate and focus on the the crucial information of the data. There are four types of autoencoders: Vanilla, Multilayer, Variational and Regularized. The Vanilla autoencoder is the simplest type, where the encoder and decoder are composed by one layer each. The Multilayer autoencoder is the extension of the Vanilla one, where encoder and decoder are composed by more than one layer. In both the previous types the network is a FCNN. The Variational autoencoder architecture is composed by a multilayer CNN, while the Regularized type is a Vanilla or Multilayer where a regularization is added to the Loss function. In our study presented in Chapter 7 we use a Multilayer autoencoder to compress the inputs and study them in a latent space of reduced dimensionality.

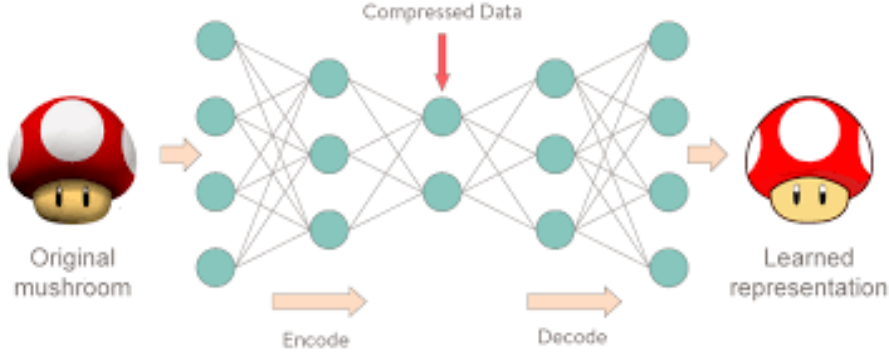


Figure 5.3: Autoencoder architecture [138].

5.3 Unsupervised Learning through K-Means

Unsupervised learning is a subset of the classification algorithms where the goal of the evaluation is to classify the inputs giving no labelled outputs during the training and leaving the algorithm to find a pattern in the data. This procedure is called clustering and the inputs are collected in the defined number of clusters. An example of unsupervised classification in three clusters is displayed in Fig. 5.4.

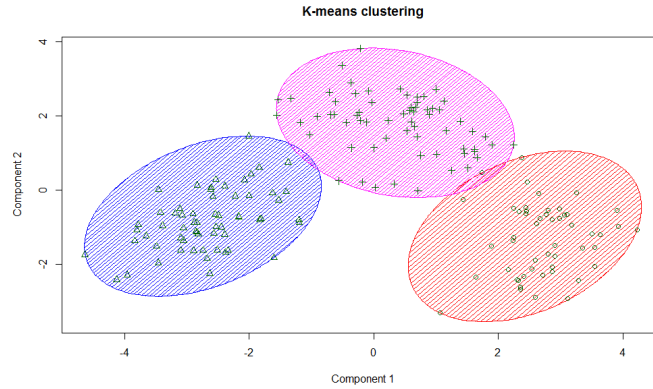


Figure 5.4: Clustering procedure example. The data set is classified in three clusters according to the K-Means criteria [138].

There are several algorithms achieving this task, in the following we describe K-Means, one of the simplest and most popular unsupervised machine learning algorithms. The first step of the K-Means clustering is to define K the number of centroids. The centroids are the points representing the centers of the clusters. They are initialized randomly, selecting K data points without replacement from the shuffled dataset. The distance of all data points from the centroids is computed and each point is assigned

to the closest cluster. The loss function of the problem is:

$$J = \sum_{i=1}^m \sum_{k=1}^K w_{ik} ||x^i - \mu_k||^2 \quad (5.2)$$

where m is the number of points in the dataset, x^i is the i^{th} point of the dataset, μ_k the k^{th} centroid and w_{ik} the weight of each point. $w_{ik}=1$ if x^i belongs to the k^{th} cluster or 0 otherwise. The loss function tells us about the quality of the clustering and its minimization is used to optimize the model in a two steps procedure where the weights are updated and the centroids redefined. The first step is to update the weights taking the derivative with respect to w_{ik} :

$$\frac{\partial J}{\partial w_{ik}} = \sum_{i=1}^m \sum_{k=1}^K ||x^i - \mu_k||^2 \rightarrow w_{ik} = \begin{cases} 1, & \text{if } k = \operatorname{argmin}_j ||x^i - \mu_j||^2 \\ 0, & \text{otherwise} \end{cases} \quad (5.3)$$

where the weight $w_{ik} = 1$ if it minimizes the Loss function, otherwise $w_{ik} = 0$.

The second step is to take the derivative with respect to μ_k :

$$\frac{\partial J}{\partial \mu_k} = 2 \sum_{i=1}^m w_{ik} (x^i - \mu_k) = 0 \rightarrow \mu_k = \frac{\sum_{i=1}^m w_{ik} x^i}{\sum_{i=1}^m w_{ik}}, \quad (5.4)$$

therefore the centroids are redefined according to the updated weights.

The procedure is repeated until there is no more variation of the centroids within a given tolerance. The inertia is an estimator of the clustering quality, representing the distance of the points within a cluster from the cluster center. In order to obtain a good model a small value of inertia is aimed for. When the dimension of the input arrays is too large K-Means can provide a bad clustering. The algorithm can suffer from the complexity of the data and misclassify some of the points. In these cases to feed the algorithm with encoded inputs can help the classification, allowing K-Means to focus on the salient features of the dataset.

The machine learning techniques described in this Chapter are adopted in our study of shape classification of Higgs boson pair invariant mass distributions presented in Chapter 7.

6 Higgs boson pair production

In this chapter we present our calculation of Higgs boson pair production at NLO QCD with full top mass dependence. The chapter is adapted from Ref. [111] where the author of this thesis is a coauthor.

Exploring the Higgs sector and the mechanism of electroweak symmetry breaking is one of the primary goals for the current and future LHC program as well as other planned experiments. While some of the properties of the Higgs boson, like its mass, spin and couplings to electroweak bosons, have been measured meanwhile impressively well [141], other parameters, like the couplings to (light) fermions, and in particular the self-coupling, are still largely unconstrained as shown in Sec. 2.4.

In the SM the strength of all Higgs boson couplings is predicted; however, effects of physics BSM may lead to deviations which, once firmly established, are a clear sign of New Physics. Since higher-order QCD corrections are known to be important in Higgs boson production processes, they need to be taken into account to improve the sensitivity to New Physics effects.

Given the energy gap between the electroweak scale at $v \simeq 250 \text{ GeV}$ and a New Physics scale Λ which could be in the TeV range, it is natural to parametrise the BSM effects in a model-independent way in an EFT framework.

Higgs boson pair production in gluon fusion is the most promising process to find out whether the Higgs boson self-coupling is Standard-Model-like. The scattering amplitude of the process is function of five anomalous couplings.

In order to quantify the different effects of the five couplings that can lead to deviations from the SM in the Higgs sector, we give results for the total NLO cross section parametrised in terms of 23 coefficients of all possible combinations of these couplings, as introduced at LO in Refs. [142, 143]. We also show differential distributions for 12 benchmark points which should be characteristic for clusters of BSM scenarios. Such clusters were identified in Refs. [143–145] at leading order and represent partitions of the BSM parameter space according to the shape of the differential distributions.

This chapter is organised as follows. Section 6.1 reviews the state of art of Higgs boson pair production. In Section 6.2, we introduce the Higgs-electroweak chiral Lagrangian applied to Higgs boson pair production at LO, including the NLO QCD corrections in Section 6.3. Section 6.4 is dedicated to the phenomenological results. We provide a parametrisation of the NLO cross section in terms of coefficients of all combinations of couplings occurring in the NLO cross section. Based on this parametrisation we show heat maps both at LO and at NLO, where we vary two couplings while keeping the others fixed to the SM values. Then we give results for total cross sections and differential distributions at twelve benchmark points and discuss their implications before we conclude.

6.1 Higgs boson pair production state of art

Early studies of Higgs boson pair production within an EFT framework can be found in Refs. [146–149]. Many phenomenological investigations about the potential of this process to reveal New Physics have been performed since, see e.g. Refs. [52, 53, 57, 142, 150–162]. In the SM, Higgs boson pair production has been calculated at leading order in Refs. [163–165]. As it is a loop-induced process, higher order calculations with full top quark mass dependence involve multi-scale two-loop integrals. Therefore, the NLO calculations until recently have been performed in the $m_t \rightarrow \infty$ limit [166], the HEFT approximation, and then rescaled by a factor B_{FT}/B_{HEFT} , B_{FT} denoting the leading order matrix element squared in the full theory. This procedure is called “Born-improved HEFT” in the following. In Refs. [167, 168], an approximation called “FT_{approx}” was introduced, which contains the full top quark mass dependence in the real radiation, while the virtual part is calculated in the HEFT approximation and rescaled at the event level by the re-weighting factor B_{FT}/B_{HEFT} .

In addition, the HEFT results at NLO and NNLO have been improved by an expansion in $1/m_t^2$ in Refs. [169–172]. The NNLO QCD corrections in the heavy-top limit have been computed in Refs. [170, 173–175], and they have been supplemented by an expansion in $1/m_t^2$ in Ref. [171] and by threshold resummation, at NLO+NNLL in Ref. [176] and at NNLO+NNLL in Ref. [177], leading to K-factors of about 1.2 relative to the Born-improved HEFT result. The N³LO QCD corrections within the HEFT approximation combined to the full NLO calculation have been computed in [178].

The full NLO corrections, including the top quark mass dependence also in the virtual two-loop amplitudes, have been calculated in Ref. [179]. Phenomenological studies at 14 TeV and 100 TeV, including variations of the Higgs boson self-coupling, have been presented in Ref. [180]. The full NLO calculation was supplemented by NLL transverse momentum resummation in Ref. [181]. It also has been matched to parton shower Monte Carlo programs [182–184], where the matched result of Ref. [182, 184] is publicly available within the POWHEG-BOX-V2 framework. The full NLO combining the exact numerical result and high-energy expansion has been calculated in Ref. [185], while a full NLO study of the uncertainty due to the scheme and scale choice for the top mass in the loops has been presented in Ref. [186]. Recent work also includes a combination of an analytic threshold expansion and a large- m_t expansion together with a Padé approximation framework [187], the expansion of the amplitudes in terms of a small Higgs transverse momentum [188], and analytic results based on a high energy expansion for the planar part of the two-loop amplitude [189].

Very recently, top quark mass effects have been incorporated in the NNLO HEFT calculation, including the full NLO result and combining one-loop double-real corrections with full top mass dependence with suitably reweighted real-virtual and double-virtual contributions evaluated in the large- m_t approximation [190].

Within a non-linear EFT framework, higher order QCD corrections have been performed in the $m_t \rightarrow \infty$ limit. The NLO QCD corrections have been calculated in Ref. [191], recently also supplemented with the case of CP-violating Higgs sectors [192]. The NNLO QCD corrections in the $m_t \rightarrow \infty$ limit including dimension 6 operators have

been presented in Ref. [193]. These calculations found rather flat K-factors, which however could be an artefact of the $m_t \rightarrow \infty$ limit. One of the main goals of this chapter is to investigate whether this feature is preserved once the full top quark mass dependence is taken into account. We calculate the NLO QCD corrections to Higgs boson pair production in gluon fusion within the non-linear EFT framework, retaining the full top quark mass dependence, based on the numerical approach developed in Ref. [179].

6.2 Non-Linear EFT for di-Higgs production

Since this process is loop-induced, at leading order both one-loop diagrams built from the LO interactions \mathcal{L}_2 , as well as tree contributions from the NLO (in the chiral expansion) Lagrangian \mathcal{L}_4 have to be taken into account. The relevant terms from the effective Lagrangian $\mathcal{L}_2 + \mathcal{L}_4$ (see Sec. 4.2) are given by [106]

$$\mathcal{L} \supset -m_t \left(c_t \frac{h}{v} + c_{tt} \frac{h^2}{v^2} \right) \bar{t} t - c_{hhh} \frac{m_h^2}{2v} h^3 + \frac{\alpha_s}{8\pi} \left(c_{ggh} \frac{h}{v} + c_{gggh} \frac{h^2}{v^2} \right) G_{\mu\nu}^a G^{a,\mu\nu}. \quad (6.1)$$

The first three couplings, c_t , c_{tt} , c_{hhh} , are from \mathcal{L}_2 , the Higgs-gluon couplings c_{ggh} and c_{gggh} from \mathcal{L}_4 [101, 112]. To lowest order in the SM $c_t = c_{hhh} = 1$ and $c_{tt} = c_{ggh} = c_{gggh} = 0$. In general, all couplings may have arbitrary values of $\mathcal{O}(1)$. Note that we have extracted a loop factor from the definition of the Higgs-gluon couplings.

The leading-order diagrams are shown in Fig. 6.1. All diagrams are at the same order

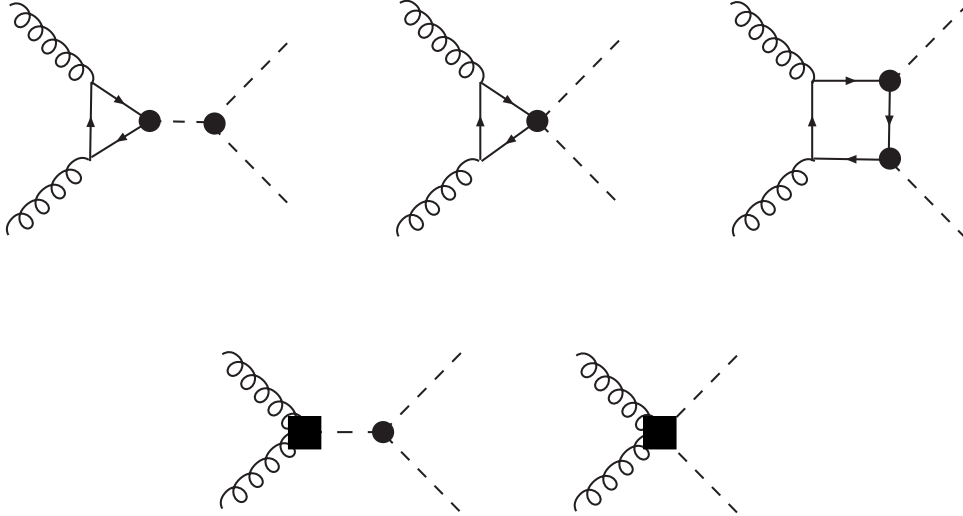


Figure 6.1: Higgs-pair production in gluon fusion at leading order in the chiral Lagrangian. The black dots indicate vertices from \mathcal{L}_2 , the black squares denote local terms from \mathcal{L}_4 .

in the chiral counting (chiral dimension 4, equivalent to one-loop order). They illustrate the interplay between leading order anomalous couplings (black dots) within loops, and

next-to-leading order terms (black squares) at tree level. All the five couplings defined in (7.1) appear in Fig. 6.1. In the following section we discuss the extension of this analysis to the next order in QCD.

6.3 Calculation of the NLO QCD corrections

Within the framework of the electroweak chiral Lagrangian, the calculation of the $gg \rightarrow hh$ amplitude can be extended to the next order in the loop expansion, that is to two-loop order, or chiral dimension 6. In full generality, this would require to also include two-loop electroweak corrections and local terms from the Lagrangian at chiral dimension 6. The latter introduce additional couplings, parametrising subleading new-physics effects. Such effects are beyond the experimental sensitivity in the foreseeable future, given that even the determination of the LO couplings in (7.1) remains a substantial challenge. On the other hand, radiative corrections from QCD are known to be very important for $gg \rightarrow hh$ and similar processes.

For this reason, we extend the calculation of $gg \rightarrow hh$ to the next order in the non-linear EFT, but restrict the NLO corrections to the effects from QCD. Within the systematics of the EWChL this approximation corresponds to including those corrections at chiral dimension 6 that come with a relative factor of the QCD coupling g_s^2 . This procedure is consistent without introducing further anomalous couplings, beyond the ones in (7.1), because this effective Lagrangian is renormalisable with respect to QCD [104]. Since the LO amplitude for $gg \rightarrow hh$ scales as $\sim g_s^2$, the NLO virtual corrections of interest to us comprise all the diagrams at two-loop order carrying a factor of g_s^4 . They exist as two-loop, one-loop and tree topologies, as illustrated in Figs. 6.2, 6.3 and 6.4, respectively.

In addition, real emission diagrams at $\mathcal{O}(g_s^3)$ have to be included as shown in Fig. 6.5.

To further clarify our approximation with respect to the full chiral expansion at NLO, we give in Fig. 6.6 a few examples of higher-order effects that are consistently neglected in our scheme:

Example (a) shows a correction from electroweak-boson exchange. It is of two-loop order, but scales as $g_s^2 g^2$, rather than g_s^4 . It is not a NLO QCD effect and we neglect it here.

Similarly, the one-loop topology in (b) counts as two-loop order, but scales only as $g_s^2 c_{hhhh}$, with c_{hhhh} the (anomalous) quartic Higgs coupling.

In example (c) we consider an anomalous top-gluon coupling of the form $Q_{tG} = y_t g_s \bar{t}_L \sigma_{\mu\nu} G^{\mu\nu} t_R$, where the top Yukawa coupling reflects the change in chirality. This operator is therefore (at least) of chiral dimension 4 (one-loop order) and the diagram in Fig. 6.6 (c) of two-loop order, but again not of order g_s^4 . Since (4.3) assumes that the top quark is weakly coupled to the (possibly strongly interacting) new-physics sector, it is more likely that the operator comes with further weak couplings from t_L and t_R and thus carries chiral dimension 6. In this case, diagram (c) is of three-loop order and

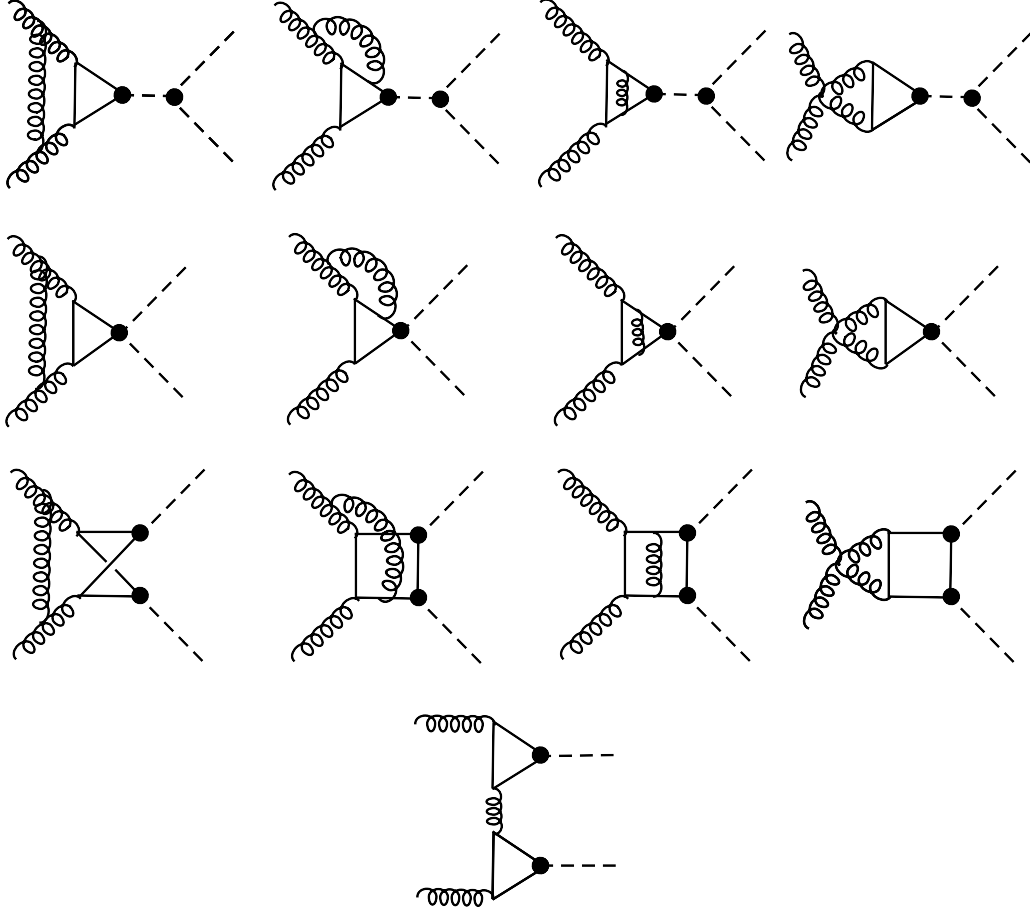


Figure 6.2: Higgs-pair production in gluon fusion at NLO: Examples for virtual two-loop diagrams at order g_s^4 .

clearly negligible. The effect of the chromomagnetic operator on single Higgs boson production has been calculated recently in the context of SMEFT in Ref. [194].

Example (d) illustrates the effect of a local Higgs-gluon interaction of chiral dimension 6, which enters at two-loop order as a tree-level topology. A possible operator would be $g_s^2 G_{\mu\nu}^a G^{a,\mu\nu} \partial_\lambda h \partial^\lambda h$. However, this effect, although of two-loop order, does not scale as g_s^4 .

Finally, we may have an operator $g_s^3 f^{abc} G_{\mu\nu}^a G^{b,\nu}_\lambda G^{c,\lambda\mu} h$, also of chiral dimension 6. Diagram (e) then amounts to a two-loop order interaction with real emission, which is beyond our approximation.

At the technical level, the NLO QCD corrections have been calculated building on the setup described in Refs. [179, 180], summarised briefly below.

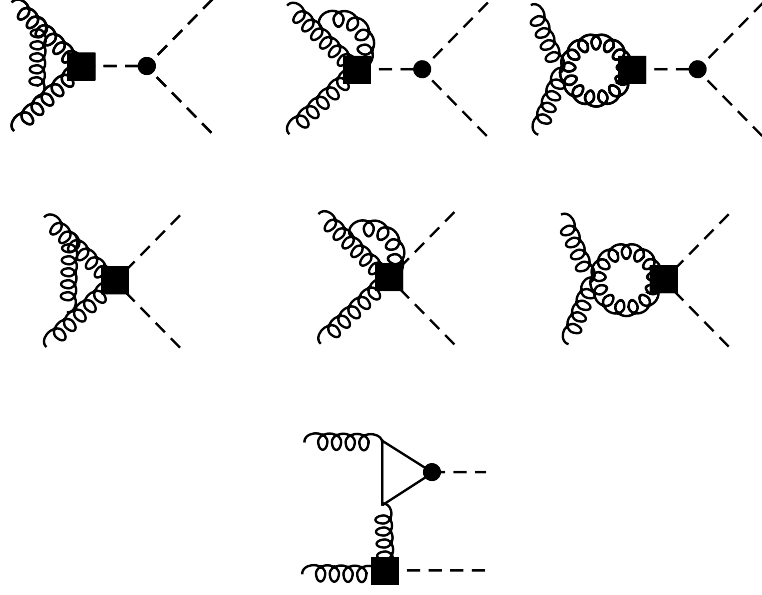


Figure 6.3: Higgs-pair production in gluon fusion at NLO: Examples for virtual one-loop diagrams at order g_s^4 .

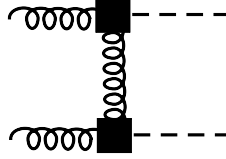


Figure 6.4: Higgs-pair production in gluon fusion at NLO: Tree diagram at order g_s^4 .

6.3.1 Virtual corrections

The virtual part of order α_s^3 consists of genuine two-loop diagrams as well as one-loop and tree-level diagrams, see Figs. 6.2, 6.3 and 6.4.

For the two-loop part, we made use of the numerical results for the two-loop virtual diagrams in the Standard Model (SM) by dividing them into two classes: diagrams containing the Higgs-boson self-coupling (“triangle-type”), and diagrams without (“box-type”). The $t\bar{t}hh$ coupling generates new two-loop topologies, see e.g. the second line of Fig. 6.2. The results for these diagrams however can be obtained from the SM triangle-type diagrams by multiplying them with the inverse Higgs boson propagator and rescaling the couplings, i.e. multiplying with $c_{t\bar{t}}/c_{hhh}$. The other two-loop diagrams occurring in our calculation have the same topologies as in the SM and therefore can be derived from the SM results by rescaling of the couplings c_t and c_{hhh} .

The one-loop part containing the Higgs-gluon contact interactions has been calculated in two ways: first, using GOSAM [195, 196] in combination with a model file in

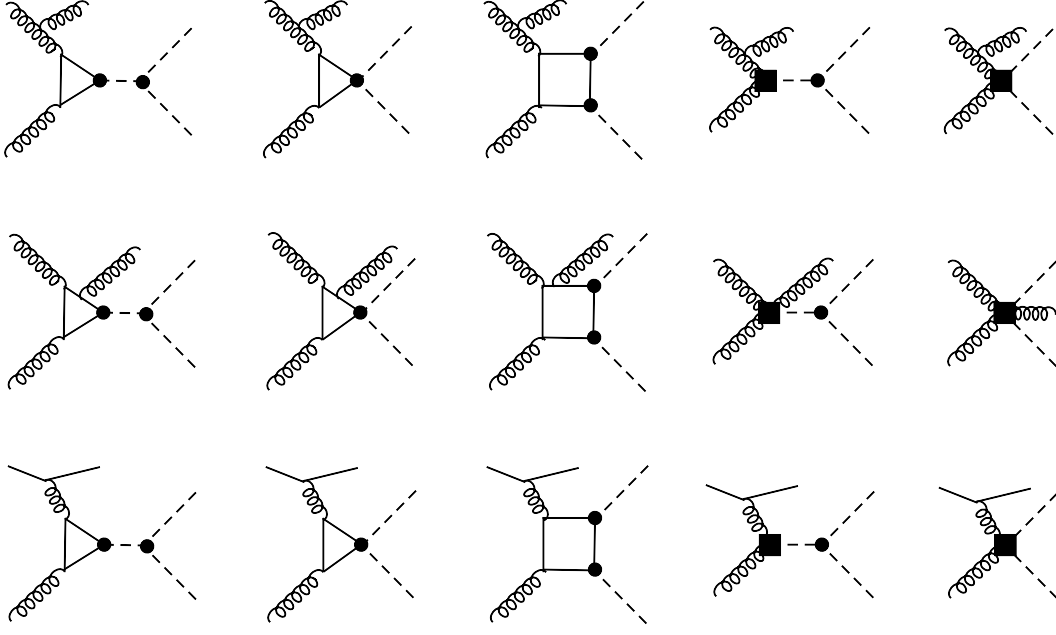


Figure 6.5: Higgs-pair production in gluon fusion at NLO: Examples for real-emission diagrams at order g_s^3 .

Universal FeynRules Output (UFO) format [197], derived from an effective Lagrangian using FEYNRULES [198], and second analytically as a cross-check. GOSAM is a software for automated one-loop calculations which is able to automatically compute one-loop QCD and/or electroweak corrections to user-defined processes. The software can be used standalone or interfaced to Monte Carlo programs via the Binoth-Les-Houches-Accord (BLHA). This tool allows to pursue phenomenological studies within and beyond the SM. The software identifies tree-level diagram as LO and one-loop diagrams as NLO, while when only the loop level is provided by the user, it recognizes the process as loop-induced. When GOSAM is used to study processes within an EFT framework, the user needs to be very carefully because of the mixing of tree level and one-loop diagrams at each order of the QCD or electroweak perturbative expansion.

As we are only considering QCD corrections, the renormalisation procedure is the same as in the SM and is described in Ref. [180].

6.3.2 Real radiation

The real corrections consist of 5-point one-loop topologies with closed top quark loops as well as tree-level diagrams, see Fig. 6.5. Both classes of diagrams have been generated with GOSAM and arranged such that interferences between the two classes are properly taken into account.

6 Higgs boson pair production

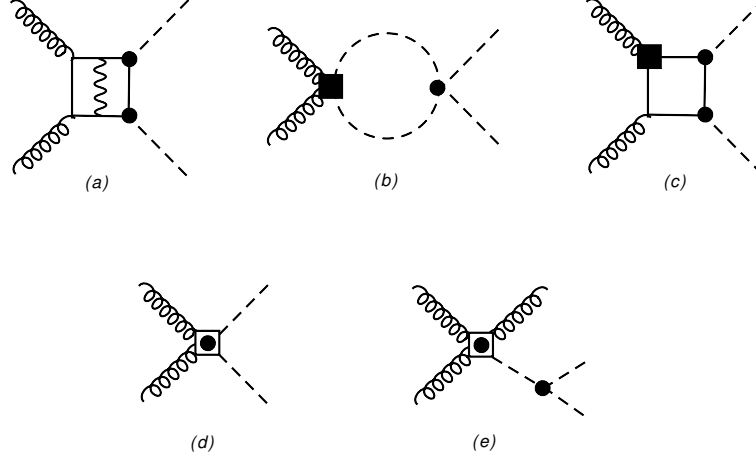


Figure 6.6: Higgs-pair production in gluon fusion at NLO: Examples for contributions that are consistently neglected within our approximation. The dotted square indicates a local term at chiral dimension 6 (two-loop order). See text for further explanation.

In order to isolate the singularities due to unresolved radiation, we use the same framework as in Ref. [180], i.e. we use the Catani-Seymour dipole formalism [84], combined with a phase space restriction parameter α as suggested in Ref. [199].

The various building blocks are assembled in a C++ program and integrated over the phase space using the VEGAS algorithm [200] as implemented in the CUBA library [201].

6.3.3 Parametrisation of the total cross section

To parametrise the deviations of the total cross section from the one in the SM, we write the LO cross section in terms of the 15 coefficients A_1, \dots, A_{15} , following Refs. [142, 143] in the following way

$$\begin{aligned} \sigma/\sigma_{SM} = & A_1 c_t^4 + A_2 c_{tt}^2 + A_3 c_t^2 c_{hhh}^2 + A_4 c_{ggh}^2 c_{hhh}^2 + A_5 c_{gghh}^2 + A_6 c_{tt} c_t^2 + A_7 c_t^3 c_{hhh} s \\ & + A_8 c_{tt} c_t c_{hhh} + A_9 c_{tt} c_{ggh} c_{hhh} + A_{10} c_{tt} c_{gghh} + A_{11} c_t^2 c_{ggh} c_{hhh} + A_{12} c_t^2 c_{gghh} \\ & + A_{13} c_t c_{hhh}^2 c_{ggh} + A_{14} c_t c_{hhh} c_{gghh} + A_{15} c_{ggh} c_{hhh} c_{gghh} . \end{aligned} \quad (6.2)$$

At NLO the coefficients A_1, \dots, A_{15} are modified and new terms appear. We find:

$$\begin{aligned} \Delta\sigma/\sigma_{SM} = & A_{16} c_t^3 c_{ggh} + A_{17} c_t c_{tt} c_{ggh} + A_{18} c_t c_{ggh}^2 c_{hhh} + A_{19} c_t c_{ggh} c_{gghh} \\ & + A_{20} c_t^2 c_{ggh}^2 + A_{21} c_{tt} c_{ggh}^2 + A_{22} c_{ggh}^3 c_{hhh} + A_{23} c_{ggh}^2 c_{gghh} . \end{aligned} \quad (6.3)$$

6.3.4 Validation of the calculation

To validate our results, we have compared the Born-improved NLO HEFT results calculated with our setup with the ones from Ref. [191], where we find agreement if we use $\mu_r = \mu_f = m_{hh} = (p_{h_1} + p_{h_2})^2$, with p_{h_1} and p_{h_2} the momenta of the two Higgs

bosons, and MSTW2008 [202] PDFs at LO/NLO for the LO/NLO calculation, along with the corresponding α_s value.¹

We also have cross-checked the results by using two independent codes, where the only common parts are the UFO model files and the SM virtual two-loop corrections.

In addition, we have compared the leading order distributions, benchmark points and fits of the coupling coefficients in the total cross section (see Eq. (6.2)) with the ones given in Refs. [142–144]. We find agreement with Ref. [142] for all A_i coefficients at the 1% level. Comparing to Refs. [143, 144], we systematically find values that differ by 15–20% for coefficients linear in c_{ggh} and by $\sim 40\%$ for the coefficient quadratic in c_{ggh} . We also compared our results with the m_{hh} distributions shown in Refs. [143, 144], finding agreement for all benchmark points except for benchmark point 8. While in Refs. [143, 144] a dip in the leading order distribution is found for benchmark point 8, we find no such dip. This is why we chose a different point of cluster 8 which does show a dip, and which we call 8a.

6.4 Phenomenological results

In this section we present numerical results for benchmark points which were identified in Ref. [143] to represent partitions of the BSM parameter space according to characteristic shapes of differential distributions, in particular the Higgs boson pair invariant mass distributions. All our results are for a centre-of-mass energy of $\sqrt{s} = 14$ TeV.

The results were computed using the PDF4LHC15_nlo_100_pdfas [73–76] parton distribution functions interfaced via LHAPDF [203], along with the corresponding value for $\alpha_s(\mu)$, with $\alpha_s(M_Z) = 0.118$. The masses of the Higgs boson and the top quark have been set to $m_h = 125$ GeV and $m_t = 173$ GeV (pole mass), respectively. The widths of the top quark (and the Higgs boson) have been set to zero. Bottom quarks are treated as massless and therefore are not included in the fermion loops. The scale uncertainties are estimated by varying the factorisation scale μ_F and the renormalisation scale μ_R around the central scale $\mu_0 = m_{hh}/2$, using the envelope of a 7-point scale variation. The latter means that we use $\mu_{R,F} = c_{R,F} \mu_0$, where $c_R, c_F \in \{2, 1, 0.5\}$, and consider each combination except the two extreme ones $c_R = 0.5, c_F = 2$ and $c_R = 2, c_F = 0.5$. In the SM case, the combinations $c_R = c_F = 0.5$ and $c_R = c_F = 2$ always coincided with the envelope of the 7 combinations to vary c_R, c_F .

¹Our default settings are to use PDF4LHC15 [76] partonic distribution functions for both the LO and the NLO results.

6.4.1 NLO cross sections and heat maps

In this subsection we will provide results for the coefficients defined in Eqs. (6.2) and (6.3), i.e. for the expression

$$\begin{aligned} \sigma^{\text{NLO}}/\sigma_{SM}^{\text{NLO}} = & A_1 c_t^4 + A_2 c_{tt}^2 + A_3 c_t^2 c_{hhh}^2 + A_4 c_{ggh}^2 c_{hhh}^2 + A_5 c_{gghh}^2 + A_6 c_{tt} c_t^2 + A_7 c_t^3 c_{hhh} \\ & + A_8 c_{tt} c_t c_{hhh} + A_9 c_{tt} c_{ggh} c_{hhh} + A_{10} c_{tt} c_{gghh} + A_{11} c_t^2 c_{ggh} c_{hhh} + A_{12} c_t^2 c_{gghh} \\ & + A_{13} c_t c_{hhh}^2 c_{ggh} + A_{14} c_t c_{hhh} c_{gghh} + A_{15} c_{ggh} c_{hhh} c_{gghh} \\ & + A_{16} c_t^3 c_{ggh} + A_{17} c_t c_{tt} c_{ggh} + A_{18} c_t c_{ggh}^2 c_{hhh} + A_{19} c_t c_{ggh} c_{gghh} \\ & + A_{20} c_t^2 c_{ggh}^2 + A_{21} c_{tt} c_{ggh}^2 + A_{22} c_{ggh}^3 c_{hhh} + A_{23} c_{ggh}^2 c_{gghh}. \end{aligned} \quad (6.4)$$

We evaluated the coefficients in two different ways: determination via projections and performing a fit, finding agreement of the results within their uncertainties. The results of the projection method, including uncertainties, are summarised in Table 6.1.

In the following we show heat maps for the ratio σ/σ_{SM} , based on the results for A_1, \dots, A_{23} . For the fixed parameters the SM values are used. Further we use $\sigma_{SM}^{\text{LO}} = 19.85 \text{ fb}$, $\sigma_{SM}^{\text{NLO}} = 32.95 \text{ fb}$.

The couplings are varied in a range which seems reasonable when taking into account the current constraints on the Higgs coupling measurements [141, 204, 205], as well as recent limits on the di-Higgs production cross section [206, 206–208].

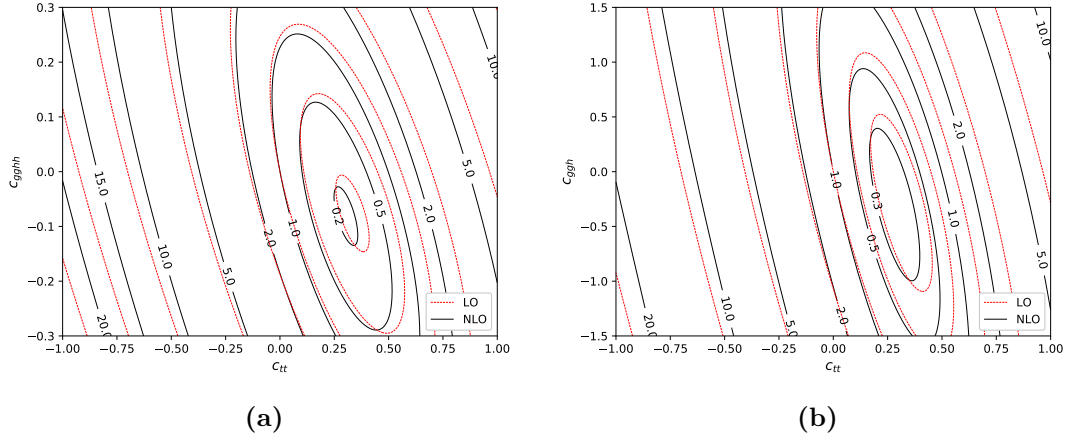


Figure 6.7: Iso-contours of σ/σ_{SM} : (a) c_{gghh} and (b) c_{ggh} versus c_{tt} .

In Fig. 6.7 we display iso-contours where the anomalous coupling c_{tt} is varied in combination with the Higgs-gluon contact interactions c_{gghh} and c_{ggh} . We show the ratio to the SM total cross section both at LO and at NLO. We can see that the NLO corrections can lead to a significant shift in the iso-contours. It also becomes apparent that the cross sections are more sensitive to variations of c_{tt} than to variations of the contact interaction c_{ggh} .

A coeff	LO value	LO uncertainty	NLO value	NLO uncertainty
A_1	2.08059	0.00163127	2.23389	0.0100989
A_2	10.2011	0.00809032	12.4598	0.0424131
A_3	0.27814	0.00187658	0.342248	0.0153637
A_4	0.314043	0.000312416	0.346822	0.00327358
A_5	12.2731	0.0101351	13.0087	0.0962361
A_6	-8.49307	0.00885261	-9.6455	0.0503776
A_7	-1.35873	0.00148022	-1.57553	0.0136033
A_8	2.80251	0.0130855	3.43849	0.0771694
A_9	2.48018	0.0127927	2.86694	0.0772341
A_{10}	14.6908	0.0311171	16.6912	0.178501
A_{11}	-1.15916	0.00307598	-1.25293	0.0291153
A_{12}	-5.51183	0.0131254	-5.81216	0.134029
A_{13}	0.560503	0.00339209	0.649714	0.0287388
A_{14}	2.47982	0.0190299	2.85933	0.193023
A_{15}	2.89431	0.0157818	3.14475	0.148658
A_{16}			-0.00816241	0.000224985
A_{17}			0.0208652	0.000398929
A_{18}			0.0168157	0.00078306
A_{19}			0.0298576	0.000829474
A_{20}			-0.0270253	0.000701919
A_{21}			0.0726921	0.0012875
A_{22}			0.0145232	0.000703893
A_{23}			0.123291	0.00650551

Table 6.1: Results for the coefficients defined in Eq. (6.4). The uncertainties are obtained from the uncertainties on the total cross sections entering the projections, using error propagation which neglects correlations between these cross sections.

Fig. 6.8 shows variations of the triple Higgs coupling c_{hhh} in combination with c_{ggh} and c_{tt} . We observe that the deviations from the SM cross section can be substantial, and again we see a rapid variation of the cross section when changing c_{tt} .

In Fig. 6.9 we display variations of c_t versus c_{hhh} , and variations of c_t versus c_{tt} . We see that values of c_t around 2.0 in combination with large negative values of c_{hhh} can enhance the cross section by two orders of magnitude. Current experimental limits suggest that the total cross section for Higgs boson pair production does not exceed about 13–24 times the SM value, assuming a SM-like shape in the distributions [206, 207]. Together with the prospects that c_t will be increasingly well constrained in the future, e.g. from measurements of $t\bar{t}H$ production [5, 30], this should allow to constrain some of the parameter space for c_{hhh} .² Fig. 6.10 shows variations of c_{gghh} versus c_{ggh}

²Note that c_t and c_{ggh} already receive indirect constraints from single Higgs boson processes, as they enter in $gg \rightarrow h$ and $h \rightarrow \gamma\gamma$.

6 Higgs boson pair production

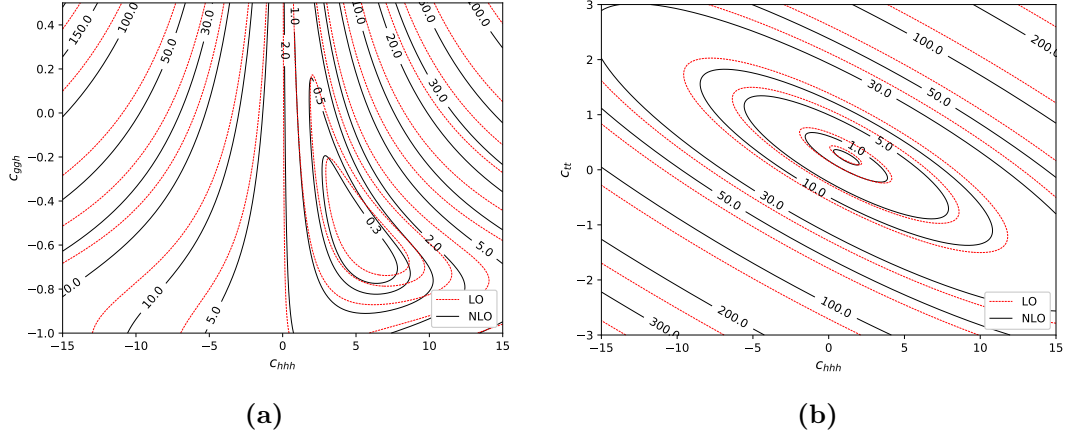


Figure 6.8: Iso-contours of σ/σ_{SM} : (a) c_{ggh} versus c_{hhh} and (b) c_{tt} versus c_{hhh} .

and c_t . We observe that the impact on the NLO corrections is milder in this case.

In Fig. 6.11 we show the K-factors, i.e. the ratio of the NLO cross section yield to the LO cross section, as a function of the coupling parameters, with the others fixed to their SM values. It shows that the rather flat K-factors which have been found [191, 193] in the $m_t \rightarrow \infty$ limit (flat with respect to variations of one of the coupling parameters) show a much stronger dependence on the coupling parameters once the full top quark mass dependence is taken into account.

6.4.2 Cross sections and distributions at several benchmark points

In the following we will show results for the benchmark points defined in Ref. [143], except for benchmark point 8, where we choose a different one (denoted as “outlier” number 5 for cluster 8 in Ref. [144]) which has a more characteristic shape, and which we call 8a. The benchmark points translated to our conventions are listed in Table 6.2.

The conventions for the definition of the couplings between our Lagrangian, given in Eq. (7.1), and the one of Ref. [143] are slightly different. In Table 6.3 we list the conversion factors to translate between the conventions.

6.4.2.1 Total cross sections

We first show the values for the total cross sections, together with their statistical uncertainties and the uncertainties from scale variations. We should point out that the cross sections for benchmark points B_3 , B_4 and B_{12} are larger than the limits measured in the $b\bar{b}\gamma\gamma$ decay channel [206, 208]. However, within the same cluster [144], i.e. the set of couplings which lead to a similar *shape* of the m_{hh} distribution, one can easily find combinations of couplings where the value of the total cross section is below the experimental exclusion bound. For example, taking the point $c_{hhh} = 1, c_t = 1, c_{tt} =$

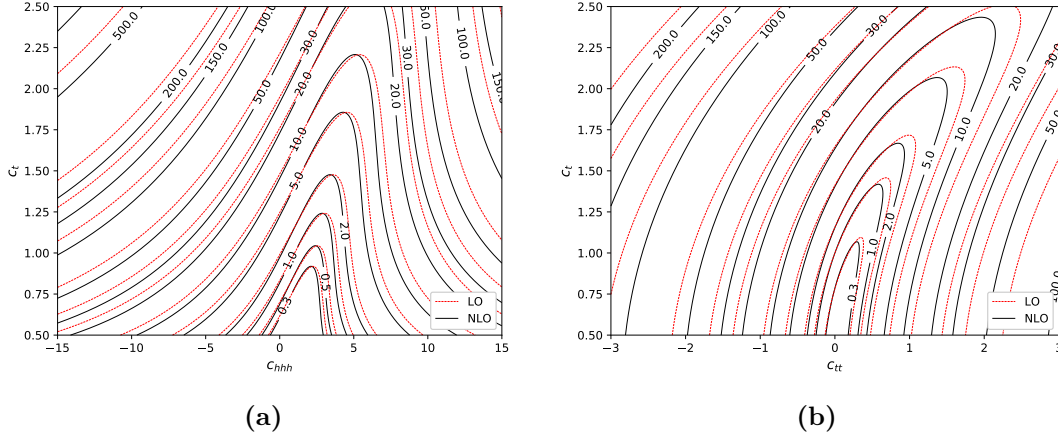


Figure 6.9: Iso-contours of σ/σ_{SM} : (a) c_t versus c_{hhh} and (b) c_t versus c_{tt} .

0, $c_{ggh} = 4/15$, $c_{gghh} = -0.2$ in cluster 4 leads to a cross section of about 1.8 times the SM cross section, still far from being excluded, see Fig. 6.16.

The large differences in the statistical uncertainties for the different benchmark points are due to the fact that the results for the virtual two-loop part are based on rescaling of the SM numerical results, which are distributed differently in the phase space. Therefore the statistical uncertainties are largest for benchmark points where the distribution in phase space is very different from the SM case. For example, benchmark 10 has a large differential cross section at low m_{hh} values, where the SM statistics is very low. This translates into the large statistical uncertainty for benchmark 10.

6.4.2.2 m_{hh} and $p_{T,h}$ distributions

Now we consider differential cross sections for the 12 benchmark points. We show the Higgs boson pair invariant mass m_{hh} distribution and the transverse momentum $p_{T,h}$ distribution of one (any) of the Higgs bosons (see Sec. 3.3). For each benchmark point we show the full NLO result in red, and compare it to the two approximations “Born-improved NLO HEFT” (purple) and FT_{approx} (green). In the FT_{approx} everything but the virtual contribution is computed in the full theory. The virtual amplitudes are treated in the same way of the “Born-improved” approximation. The leading order (yellow) as well as the SM results are also shown (blue NLO, black LO). The lower ratio plot shows the ratio of the two approximate results to the full NLO result. The upper ratio plot shows the differential BSM K-factor, i.e. NLO_{BSM}/LO_{BSM} , both evaluated with the same PDFs.

Fig. 6.12 corresponds to a benchmark point with no Higgs-gluon contact interactions, but an enhanced triple Higgs coupling and a nonzero $t\bar{t}hh$ interaction with $c_{tt} < 0$. The total cross section is about 6 times the SM cross section, and the shape of the m_{hh} distribution is completely different from the SM. In fact, one can show analytically that the LO cross section in the $m_t \rightarrow \infty$ limit exactly vanishes near $m_{hh} = 364$ GeV, which

6 Higgs boson pair production

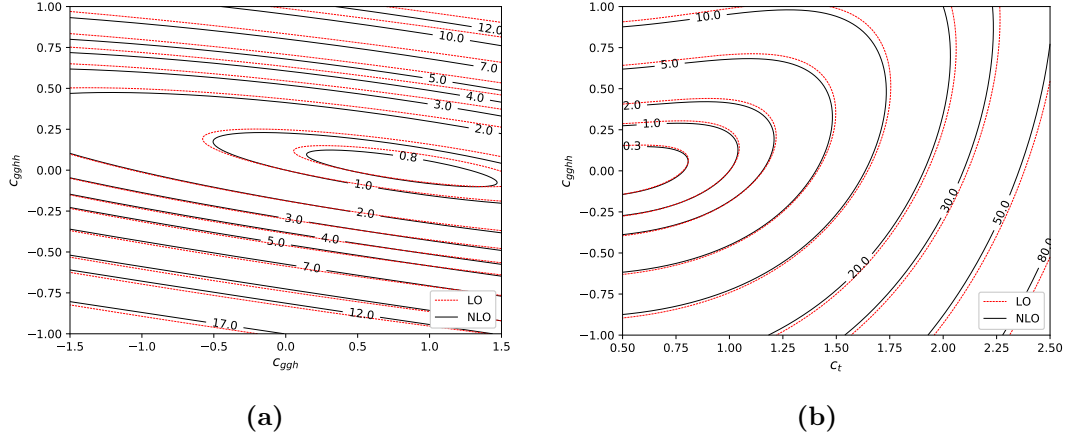


Figure 6.10: Iso-contours of σ/σ_{SM} : (a) c_{gghh} versus c_{ggh} and (b) c_{gghh} versus c_t .

relates to the dip in the distribution. The huge enhancement at low m_{hh} values is due to the large value of c_{hhh} .

Fig. 6.13, corresponding to benchmark 2, shows a very different behaviour. The result is very much suppressed in the region where the SM shows a peak, while there is a large enhancement in the tail of both the m_{hh} and the $p_{T,h}$ distributions. The enhancement in the tail is mainly due to the nonzero c_{gghh} value, as the amplitude proportional to c_{gghh} grows like \hat{s} [142]. We also notice that the approximations “Born-improved NLO HEFT” and FT_{approx} cannot describe the pattern around the $2m_t$ threshold, where the nonzero value of c_{tt} seems to play a significant role. The K-factor for benchmark 2 is very non-homogeneous around the dip in the m_{hh} distribution, and can reach up to a factor of three. This is a clear example where rescaling the LO result with a K-factor obtained from higher order calculations in the HEFT approximation would lead to very different results.

Benchmark point 3, shown in Fig. 6.14, has the same values for c_{hhh} and c_t as benchmark point 2 (the SM values), but the distributions show a very different behaviour. As in the SM, there is a peak around the $2m_t$ threshold, but the cross section is largely enhanced, not only in the peak region. As mentioned above, with a total cross section of about 32 times the SM NLO cross section, this parameter point is above the current limit deduced at 95% CL from the measured $pp \rightarrow HH \rightarrow \gamma\gamma b\bar{b}$ cross section [206, 208].

Benchmark point 4, shown in Fig. 6.15, has negative values for c_{hhh} and c_{tt} , a slightly increased Yukawa coupling c_t , and no Higgs-gluon contact interactions. This combination removes the destructive interference between different types of diagrams present in the SM, and therefore leads to a very large cross section. The differential K-factor is about 2, as for the other benchmarks, and rather constant over the whole m_{hh} range (whereas for benchmark 2, the differential K-factor is far from being homogeneous). Benchmark 4 is the one with the largest cross section of all the considered benchmark points, with a total cross section of about 270 times the SM one. This point in param-

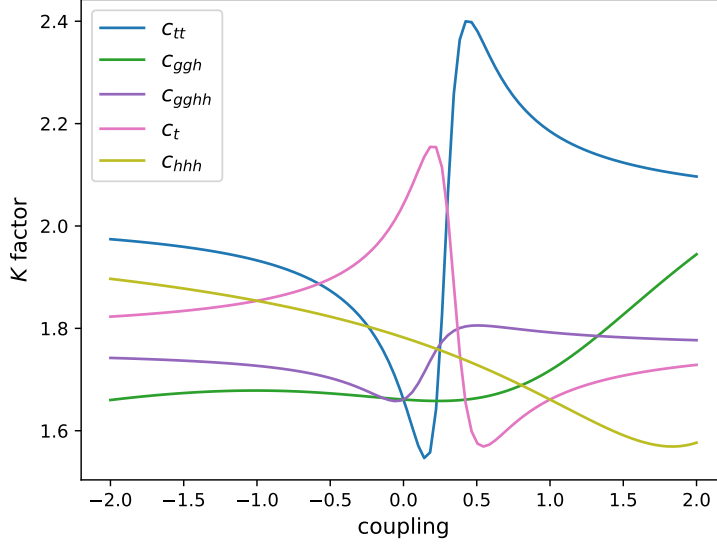


Figure 6.11: K-factors for the total cross section as a function of the different couplings.

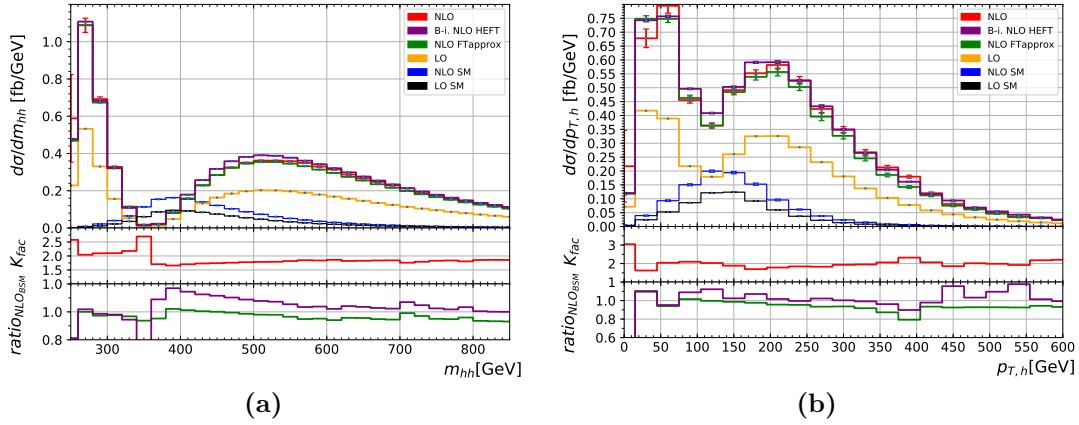


Figure 6.12: Higgs boson pair invariant mass distribution and transverse momentum distribution of one of the Higgs bosons for benchmark point 1, $c_{hhh} = 7.5$, $c_t = 1$, $c_{tt} = -1$, $c_{ggh} = c_{gggh} = 0$. The ratio plot with the K-factor shows $\text{NLO}_{\text{BSM}}/\text{LO}_{\text{BSM}}$. The lower ratio plot shows the ratios (Born-improved NLO HEFT)/ NLO_{BSM} (purple) and $\text{FT}_{\text{approx}}/\text{NLO}_{\text{BSM}}$ (green).

eter space is excluded already. Therefore, in Fig. 6.16, we also show results for another point from cluster 4, defined by $c_{hhh} = 1$, $c_t = 1$, $c_{tt} = 0$, $c_{ggh} = 4/15$, $c_{gggh} = -0.2$, which leads to a similar shape as benchmark point 4, but to $\sigma/\sigma_{SM} = 1.8$, and hence is not yet excluded. This parameter point also has the interesting feature that the dis-

6 Higgs boson pair production

Benchmark	c_{hhh}	c_t	c_{tt}	c_{ggh}	c_{gghh}
1	7.5	1.0	-1.0	0.0	0.0
2	1.0	1.0	0.5	$-\frac{1.6}{3}$	-0.2
3	1.0	1.0	-1.5	0.0	$\frac{0.8}{3}$
4	-3.5	1.5	-3.0	0.0	0.0
5	1.0	1.0	0.0	$\frac{1.6}{3}$	$\frac{1.0}{3}$
6	2.4	1.0	0.0	$\frac{0.4}{3}$	$\frac{0.2}{3}$
7	5.0	1.0	0.0	$\frac{0.4}{3}$	$\frac{0.2}{3}$
8a	1.0	1.0	0.5	$\frac{0.8}{3}$	0.0
9	1.0	1.0	1.0	-0.4	-0.2
10	10.0	1.5	-1.0	0.0	0.0
11	2.4	1.0	0.0	$\frac{2.0}{3}$	$\frac{1.0}{3}$
12	15.0	1.0	1.0	0.0	0.0
SM	1.0	1.0	0.0	0.0	0.0

Table 6.2: Benchmark points used for the distributions shown below.

EWChL Eq. (7.1)	Ref. [143]
c_{hhh}	κ_λ
c_t	κ_t
c_{tt}	c_2
c_{ggh}	$\frac{2}{3}c_g$
c_{gghh}	$-\frac{1}{3}c_{2g}$

Table 6.3: Translation between the conventions for the definition of the anomalous couplings.

tributions for NLO SM and LO BSM almost coincide. However, there is no degeneracy with the SM distribution once the NLO corrections are taken into account.

Fig. 6.17 shows distributions for benchmark point 5, where c_{tt} is zero and c_{hhh} and c_t are as in the SM, while the Higgs-gluon interactions are nonzero. Similar to benchmark point 2, we observe a dip near $m_{hh} = 350$ GeV, but the LO HEFT amplitude does not vanish there. The total cross section for benchmark point 5 is very similar to the SM one. This is an example where differential measurements are crucial to establish a clear BSM signal. The $p_{T,h}$ distribution shows the rather unexpected behaviour that $\text{FT}_{\text{approx}}$ and Born-improved HEFT drop very rapidly at large values of $p_{T,h}$. The reason is that the rescaling factor B_{FT}/B_{HEFT} becomes very large as the energy increases, because B_{HEFT} does not grow with \hat{s} for this combination of couplings, but becomes very small. Therefore the negative virtual corrections are multiplied by a very large factor, leading to the fall-off of the green and purple curves in the tail of the $p_{T,h}$ distribution.

Benchmark point 6, shown in Fig. 6.18, also shows a dip, related to the fact that the LO HEFT amplitude exactly vanishes at $m_{hh} = 429$ GeV. In addition it has a large

Benchmark	σ_{NLO} [fb]	K-factor	scale uncert. [%]	stat. uncert. [%]	$\frac{\sigma_{NLO}}{\sigma_{NLO,SM}}$
B_1	194.89	1.88	$^{+19}_{-15}$	1.6	5.915
B_2	14.55	1.88	$^{+5}_{-13}$	0.56	0.4416
B_3	1047.37	1.98	$^{+21}_{-16}$	0.15	31.79
B_4	8922.75	1.98	$^{+19}_{-16}$	0.39	270.8
B_5	59.325	1.83	$^{+4}_{-15}$	0.36	1.801
B_6	24.69	1.89	$^{+2}_{-11}$	2.1	0.7495
B_7	169.41	2.07	$^{+9}_{-12}$	2.2	5.142
B_{8a}	41.70	2.34	$^{+6}_{-9}$	0.63	1.266
B_9	146.00	2.30	$^{+22}_{-16}$	0.31	4.431
B_{10}	575.86	2.00	$^{+17}_{-14}$	3.2	17.48
B_{11}	174.70	1.92	$^{+24}_{-8}$	1.2	5.303
B_{12}	3618.53	2.07	$^{+16}_{-15}$	1.2	109.83
SM	32.95	1.66	$^{+14}_{-13}$	0.1	1

Table 6.4: Total cross sections at NLO (second column) including the K-factor (third column), scale uncertainties (4th column) and statistical uncertainties (5th column) and the ratio to the SM total NLO cross section (6th column).

enhancement of the low m_{hh} region due to the value $c_{hhh} = 2.4$. Note that this value for c_{hhh} is very close to the point where the total cross section as a function of c_{hhh} goes through a minimum if all other couplings are kept SM-like.

Benchmark point 7, shown in Fig. 6.19, has the same values for c_{ggh} , c_{gghh} , c_t and c_{tt} as benchmark point 6, but a different value for c_{hhh} ($c_{hhh} = 5$). This makes the dip disappear completely, leading to a total cross section which is about 6.7 times larger than the one for benchmark 6, and a large enhancement of the low m_{hh} and low $p_{T,h}$ regions. The distributions also show that the full top quark mass dependence in the “triangle-type” diagrams containing c_{hhh} , which dominate the low m_{hh} region, seems to play a significant role, as the full NLO result is quite different from the approximate results.

Benchmark point 8a, displayed in Fig. 6.20, again shows a characteristic dip just before the $2m_t$ threshold. It is also an example where the total cross section is very similar to the SM one, but the shape of both the m_{hh} and the $p_{T,h}$ distributions clearly discriminates the SM from the BSM case.

Benchmark point 9, displayed in Fig. 6.21, shows a large enhancement in the tails of the distributions, similar to benchmarks 2 and 3, which can be attributed mainly to the rather large value of c_{gghh} , in combination with a non-zero value of c_{tt} .

For benchmark point 10, shown in Fig. 6.22, the large value of $c_{hhh} = 10$ completely dominates the shape, leading to a large enhancement in the low m_{hh} and $p_{T,h}$ regions. With a value for the total cross section which is about 17 times larger than the SM

6 Higgs boson pair production

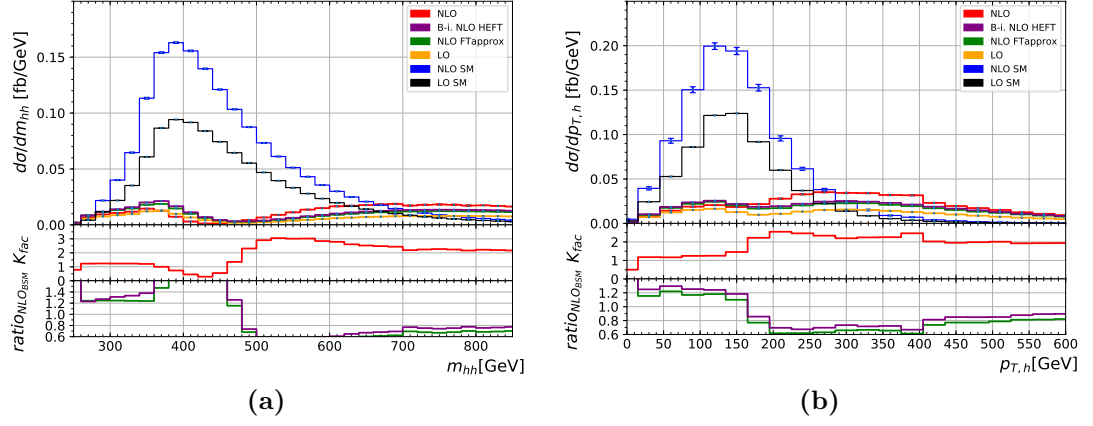


Figure 6.13: Same as Fig. 6.12 but for benchmark point 2, $c_{hhh} = 1, c_t = 1, c_{tt} = 0.5, c_{ggh} = -8/15, c_{gghh} = -0.2$.

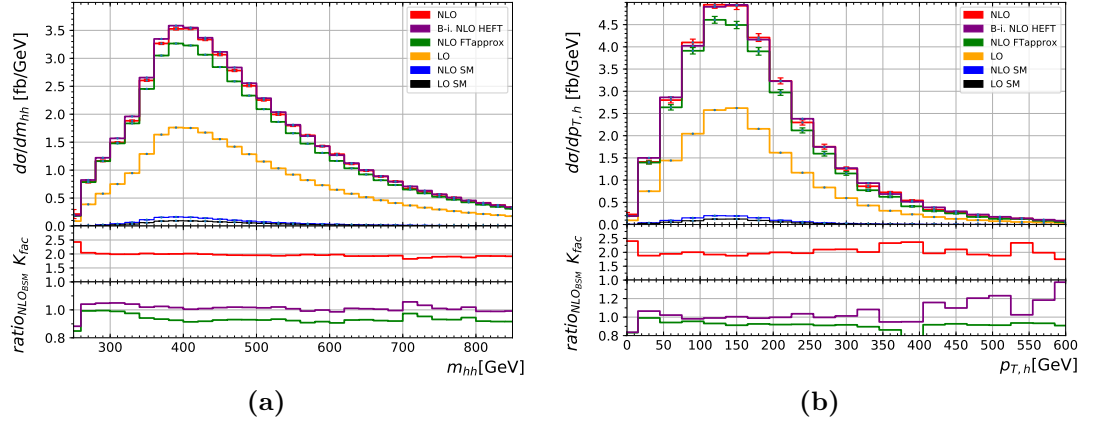


Figure 6.14: Same as Fig. 6.12 but for benchmark point 3, $c_{hhh} = 1, c_t = 1, c_{tt} = -1.5, c_{ggh} = 0, c_{gghh} = 4/15$.

cross section, benchmark point 10 is still allowed by the limits given by CMS [206], where separate limits for the various benchmark points are given.

Benchmark point 11, displayed in Fig. 6.23, has the same value for c_{hhh} as benchmark 6, which is the one where the destructive interference would be maximal if all other couplings are kept SM-like. However, the destructive interference is compensated by the large and non-zero values of c_{ggh} and c_{gghh} , such that the total cross section for benchmark 11 is about 5 times larger than the SM cross section. In view of the fact that this benchmark point is dominated by the Higgs-gluon contact interactions parametrised by c_{ggh} and c_{gghh} , it is not a surprise that the approximations $\text{FT}_{\text{approx}}$ and Born-improved HEFT agree quite well with the full calculation, as all three curves have these contributions in common, while the part which differs is damped by the destructive interference.

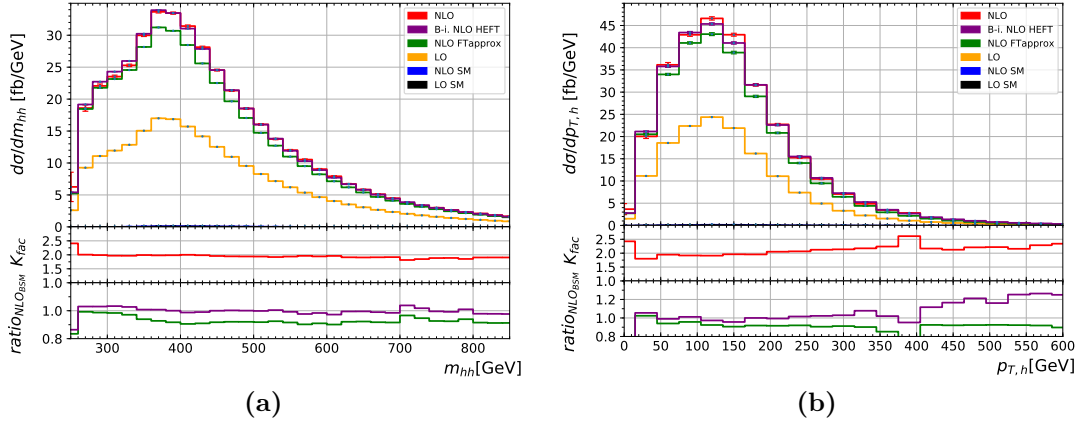


Figure 6.15: Same as Fig. 6.12 but for benchmark point 4, $c_{hhh} = -3.5, c_t = 1.5, c_{tt} = -3, c_{ggh} = c_{gghh} = 0$.

Benchmark point 12, shown in Fig. 6.24, has all couplings SM-like except $c_{tt} = 1$ and c_{hhh} , where for the latter an extreme value of $c_{hhh} = 15$ is chosen, leading to a cross section about 100 times larger than the SM cross section. This scenario is already ruled out by current LHC measurements.

All the distributions show that the NLO K-factors are large, being about a factor of two or larger. Therefore it is essential to take NLO corrections into account. The approximations where the top quark mass dependence is only partly taken into account also differ substantially in the shape from the full result for some of the benchmark points, which emphasises the importance of including the full top quark mass dependence.

In Fig. 6.25, we show the LO and NLO scale variation bands for benchmark point 5. This benchmark point is an example where the scale variation band of the 7-point scale variation mainly decreases the differential cross section over almost the whole m_{hh} range, where the upper limit of the scale variation band is mostly given by the combination $\mu_F = \mu_0/2, \mu_R = \mu_0$, for some of the bins also by $\mu_F = \mu_0, \mu_R = 2\mu_0$. In the SM, the upper limit of the 7-point scale variation band is given by $\mu_F = \mu_R = \mu_0$ for all bins of the m_{hh} distribution. We further notice that LO and NLO scale variation bands do not overlap for the m_{hh} distribution. However, this feature is also present in the SM.

6.4.2.3 Discussion of the benchmark points

Attempting a more global view of the behaviour of the m_{hh} distribution as a function of the five BSM parameters, we can identify the following patterns: a dip in the m_{hh} distributions is present for benchmark points 1, 2, 5, 6 and 8a. The presence of a non-zero value for c_{tt} or c_{ggh} is a characteristic feature of many parameter space points that show a dip in the m_{hh} distribution, but this is not a necessary condition for the presence of the dip. For instance, points with $c_{hhh} \simeq 2.5c_t$ and the other couplings

6 Higgs boson pair production

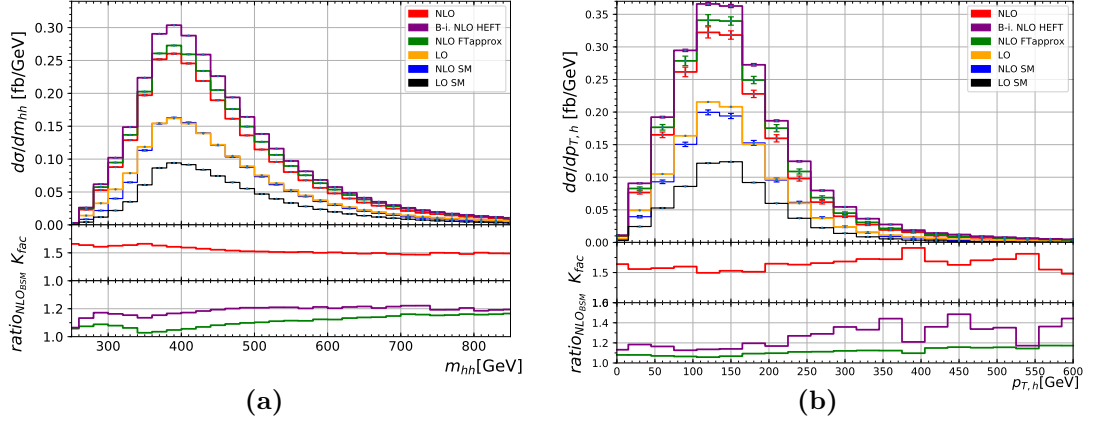


Figure 6.16: A point from cluster 4, $c_{hhh} = 1, c_t = 1, c_{tt} = 0, c_{ggh} = 4/15, c_{gghh} = -0.2$, which leads to a similar shape as benchmark point 4, but to a much smaller cross section.

vanishing also show such a dip. For the subset (1, 2, 6) of the above points there is a m_{hh} value where the LO amplitude in the $m_t \rightarrow \infty$ limit exactly vanishes, which is a feature that can cause the dip. The low m_{hh} region is enhanced for benchmark points 1, 6, 7, 10, 11, 12, which is mainly due to the large value of c_{hhh} , as the matrix element squared proportional to c_{hhh}^2 for large \hat{s} behaves like $m_h^2/\hat{s} \log^2(m_t^2/\hat{s})$ [142] and therefore dominates at low values of \hat{s} . The term proportional to c_{tt}^2 for large \hat{s} behaves like $\log^2(m_t^2/\hat{s})$ and seems to partially cancel the logarithmic terms from c_{hhh} , such that benchmark 4 has a SM-like shape even though the absolute value for c_{hhh} is large. The matrix element squared proportional to c_{gghh} grows like \hat{s} , this is why for benchmark points which have large values of c_{gghh} , the tail of the m_{hh} distribution is enhanced.

In order to assess the effect of a variation of c_{tt} while the other couplings are fixed to their SM values, we show m_{hh} distributions for the c_{tt} values $c_{tt} = -2, -1, 0, 1, 2$ in Fig. 6.26. The minimum of the cross section is at $c_{tt} \sim 0.25$. We observe that the enhancement of the cross section as $|c_{tt}|$ increases is growing more rapidly for negative values of c_{tt} , see also Fig. 6.8b. The shape changes compared to the SM are most pronounced in the low m_{hh} region.

6.5 Conclusions

We have calculated the NLO QCD corrections with full m_t dependence to Higgs boson pair production within the framework of the electroweak chiral Lagrangian, a non-linearly realised Effective Field Theory in the Higgs sector, which allows to focus on anomalous Higgs boson properties. This restricts the BSM parameter space to five possibly anomalous couplings, $c_{hhh}, c_t, c_{tt}, c_{ggh}$ and c_{gghh} .

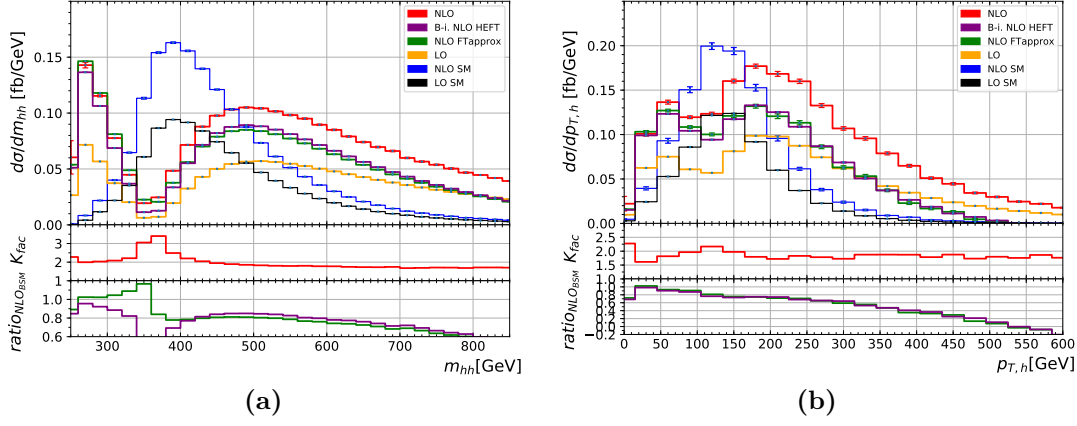


Figure 6.17: Same as Fig. 6.12 but for benchmark point 5, $c_{hhh} = 1, c_t = 1, c_{tt} = 0, c_{ggh} = 8/15, c_{gghh} = 1/3$.

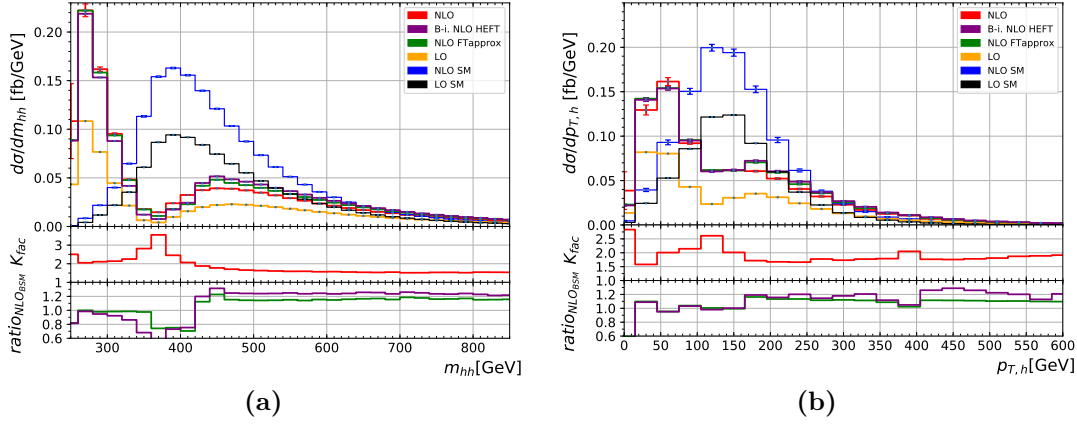


Figure 6.18: Same as Fig. 6.12 but for benchmark point 6, $c_{hhh} = 2.4, c_t = 1, c_{tt} = 0, c_{ggh} = 2/15, c_{gghh} = 1/15$.

We gave a parametrisation of the total NLO cross section and of the m_{hh} distribution in terms of 23 coefficients of all combinations of these couplings, and also showed iso-contours of LO and NLO cross section ratios σ/σ_{SM} for two-dimensional projections of the parameter space. These studies showed that the cross sections are very sensitive to variations of c_{tt} , the effective $t\bar{t}hh$ coupling, and that the K-factors can be large and non-uniform as the anomalous couplings are varied.

We have also shown differential cross sections for m_{hh} and $p_{T,h}$ at several benchmark points which exhibit characteristic shapes of the distributions. The differential K-factors for the m_{hh} distributions are of the order of two, but can reach up to three and can be very non-uniform over the m_{hh} range. This means that a rescaling of the LO distribution with a global K-factor can be rather misleading.

6 Higgs boson pair production

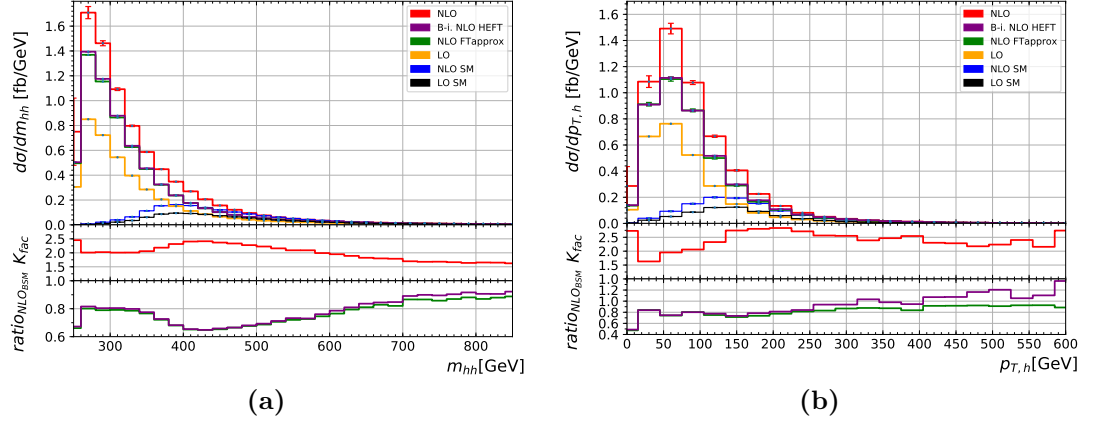


Figure 6.19: Same as Fig. 6.12 but for benchmark point 7, $c_{hhh} = 5, c_t = 1, c_{tt} = 0, c_{ggh} = 2/15, c_{gghh} = 1/15$.

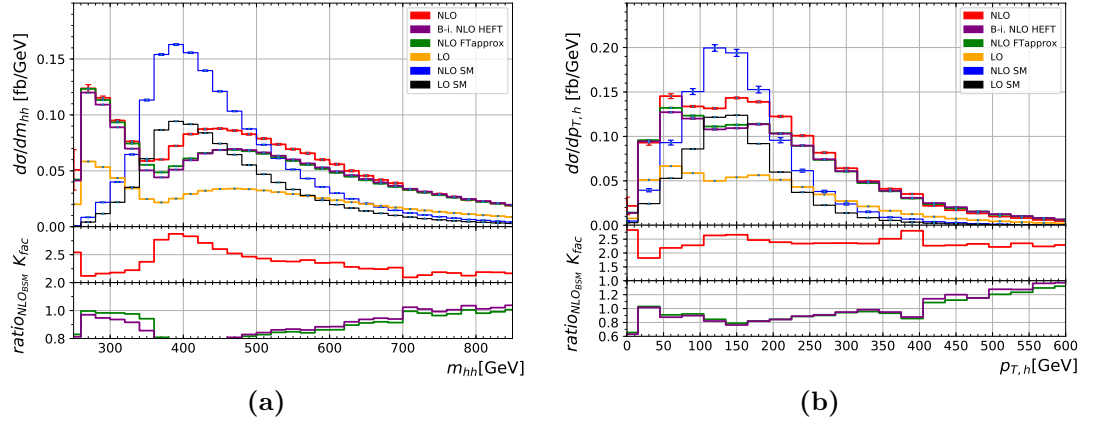


Figure 6.20: Same as Fig. 6.12 but for benchmark point 8a, $c_{hhh} = 1, c_t = 1, c_{tt} = 0.5, c_{ggh} = 4/15, c_{gghh} = 0$.

Some combinations of couplings lead to a huge enhancement of the cross section, others lead to a total cross section which is nearly degenerate to the SM one, but the corresponding m_{hh} and $p_{T,h}$ distributions have a shape which is very different from the SM one, and therefore should have discriminating power even with low statistics, which emphasises the importance of measuring distributions.

Our analytical parametrisation of the total NLO cross section and of the m_{hh} distribution in terms of all possible combinations of anomalous couplings should open the door to further studies of the considered BSM parameter space and lead to refined limits on anomalous Higgs boson couplings in the not too distant future.

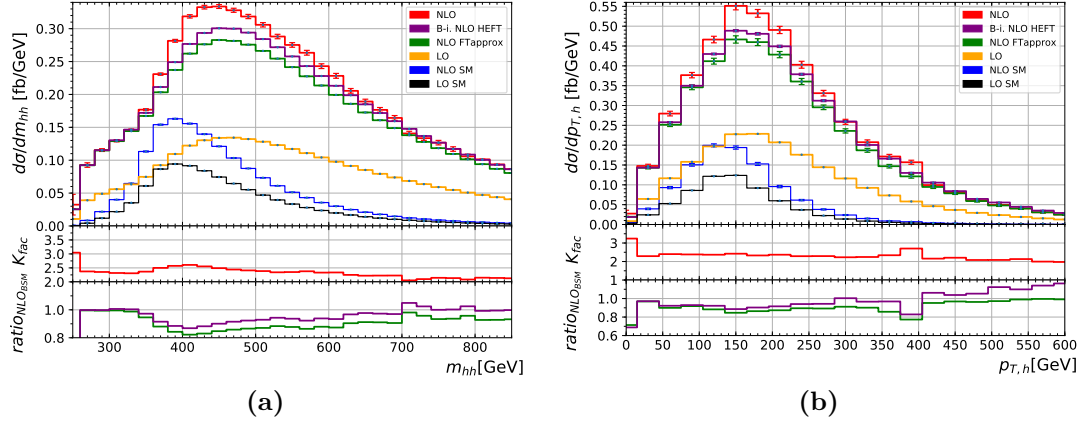


Figure 6.21: Same as Fig. 6.12 but for benchmark point 9, $c_{hhh} = 1, c_t = 1, c_{tt} = 1, c_{ggh} = -0.4, c_{gghh} = -0.2$.

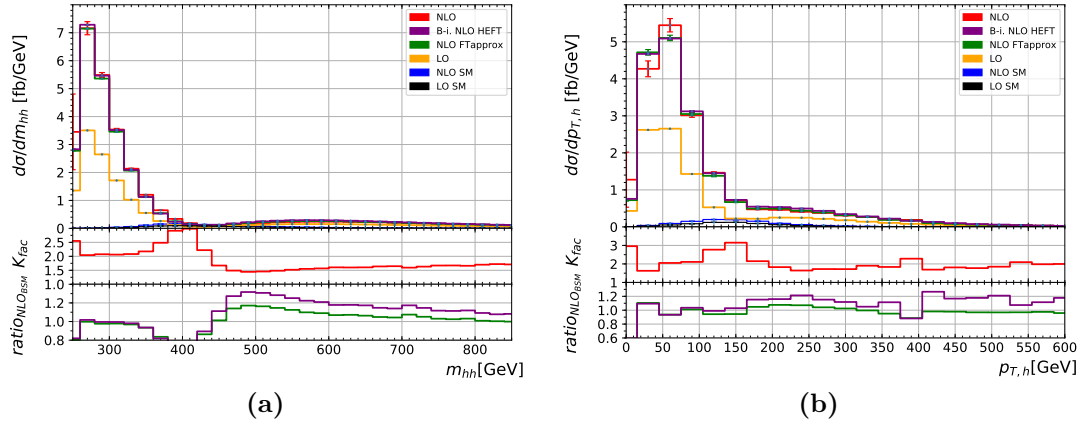


Figure 6.22: Same as Fig. 6.12 but for benchmark point 10, $c_{hhh} = 10, c_t = 1.5, c_{tt} = -1, c_{ggh} = c_{gghh} = 0$.

6 Higgs boson pair production

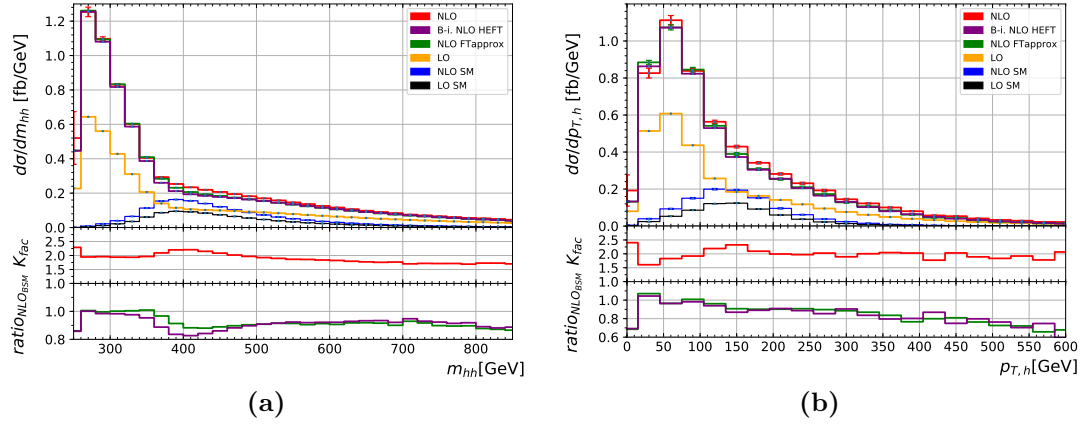


Figure 6.23: Same as Fig. 6.12 but for benchmark point 11, $c_{hhh} = 2.4, c_t = 1, c_{tt} = 0, c_{ggh} = 2/3, c_{gghh} = 1/3$.

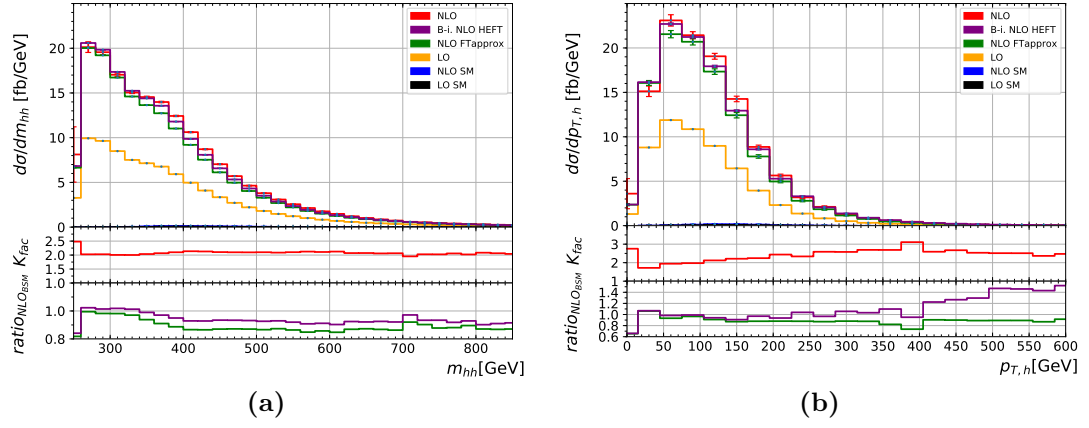


Figure 6.24: Same as Fig. 6.12 but for benchmark point 12, $c_{hhh} = 15, c_t = 1, c_{tt} = 1, c_{ggh} = c_{gghh} = 0$.

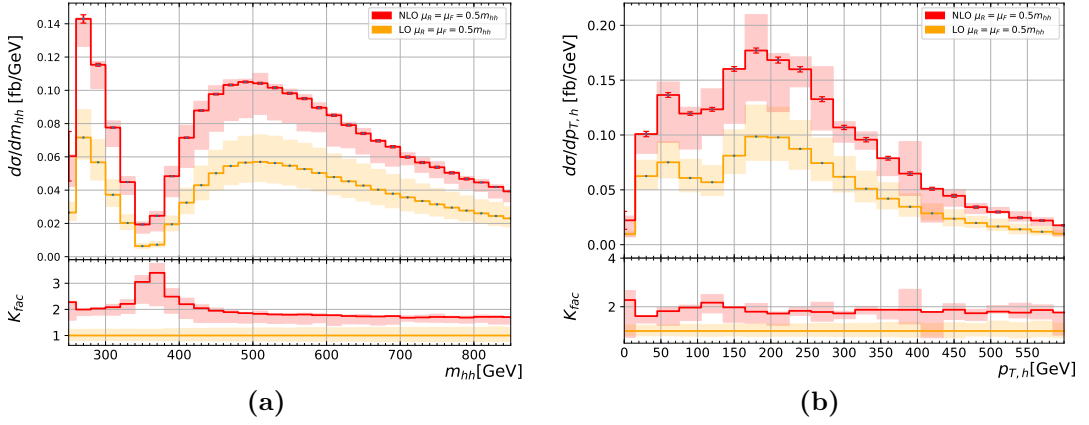


Figure 6.25: Scale variations for benchmark point 5.

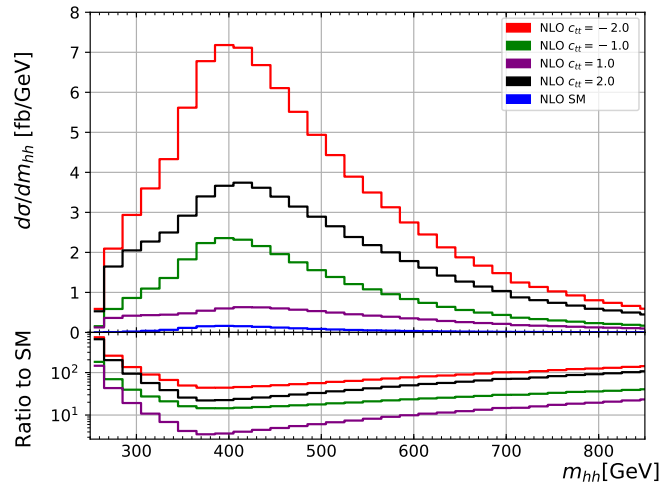


Figure 6.26: Higgs boson pair invariant mass distributions for various values of c_{tt} .

7 Higgs boson pair production and m_{hh} distribution shape analysis

We have shown in the previous Chapter that even small anomalous couplings can lead to characteristic shape changes of the invariant mass distribution m_{hh} of Higgs boson pair production. Therefore it is important to investigate in which way the shapes are influenced by a certain configuration in the coupling parameter space. The chapter is adapted from [31] where the author of this thesis is a coauthor.

The idea of a shape analysis has been pursued already in various ways based on LO studies, see e.g. Refs. [142–145, 209, 210]. In Ref. [143], a cluster analysis is proposed to define 12 benchmark points (presented in Chapter 6) in a 5-dimensional non-linear EFT parameter space, which result from clusters of “similar” shapes. The similarity measure in this case is based on a binned likelihood ratio using LO predictions for the observables m_{hh} , $p_{T,h}$ and the cosine of the polar angle of one Higgs boson with respect to the beam axis $\cos\theta^*$.

As a function of the 5-dimensional coupling parameter space, the m_{hh} distribution can have a few characterising features, such as an enhanced low- m_{hh} region, a double peak, a single peak or an enhanced tail. Some of these features can be attributed rather easily to a certain anomalous coupling. For example, an enhanced low- m_{hh} region is naturally produced by values of $|c_{hhh}| > 3$, as this leads to a dominance of the triangle-type contributions, which are suppressed by $1/\hat{s}$ and therefore die out quickly for larger m_{hh} values. Other features of the m_{hh} -shape, like a double peak or a SM-like distribution, are harder to attribute to a certain coupling configuration, as there are a multitude of configurations leading to such shapes. This is also reflected in the cluster analysis proposed in Ref. [144], where (a) very different coupling configurations can end up in the same cluster, and (b) the same cluster can contain shapes which “by eye” look quite different (for example “double peak” and “single peak”).

Therefore it is desirable to seek for alternative methods to extract information about the underlying parameter space from the shape of distributions in Higgs boson pair production. In this Chapter we first classify the shapes of Higgs boson pair invariant mass distributions, calculated at full NLO, into four characteristic types. We visualise the underlying 5-dimensional EFT parameter space producing these shape types, projecting onto 2-dimensional subspaces. We also comment on the shape of the $p_{T,h}$ distribution. Then we refine the shape analysis, applying an unsupervised learning algorithm based on an autoencoder to identify patterns in the shapes of the m_{hh} distribution. We use the **KMeans** clustering algorithm from **scikit-learn** [211] and ask for a classification of the shapes into four to eight clusters. The unsupervised classification into four clusters is compared to the analysis based on predefined clusters for

validation. Then the number of clusters is increased with the aim to find an “optimal” number of clusters, in the sense that it captures distinct shape features, but does not focus on minor details. One aim of this study is to offer an alternative to the cluster analysis proposed in Refs. [143–145] and earlier work. While the benchmark points representing a cluster are isolated points in the parameter space, the procedure we propose here allows us to associate certain shapes more straightforwardly with distinct regions in the parameter space.

Shape analysis with machine learning has been applied already to constrain anomalous Higgs-vector boson couplings in HZ production [212].

This chapter is structured as follows: In Section 7.1 we explain the framework our data samples are based on. We define four different shape types for the m_{hh} distribution and visualise the parameter space underlying the predefined shape types. In Section 7.2.1 we describe our cluster analysis based on unsupervised learning and show how the clusters found by this procedure relate to the underlying parameter space, before we conclude.

7.1 Classification through predefined shape types

As a starting point we use the effective Lagrangian in the EWChL framework relevant for Higgs boson pair production [112] already presented in Chapter 6:

$$\mathcal{L} \supset -m_t \left(c_t \frac{h}{v} + c_{tt} \frac{h^2}{v^2} \right) \bar{t} t - c_{hhh} \frac{m_h^2}{2v} h^3 + \frac{\alpha_s}{8\pi} \left(c_{ggh} \frac{h}{v} + c_{gghh} \frac{h^2}{v^2} \right) G_{\mu\nu}^a G^{a,\mu\nu}. \quad (7.1)$$

In the SM $c_t = c_{hhh} = 1$ and $c_{tt} = c_{ggh} = c_{gghh} = 0$.

We produce our data using the differential NLO m_{hh} parametrised in terms of coefficients A_i for each coupling combination occurring in the (differential) NLO cross section, which allows for a fast evaluation:

$$\begin{aligned} \frac{d\sigma}{dm_{hh}} = & A_1 c_t^4 + A_2 c_{tt}^2 + A_3 c_t^2 c_{hhh}^2 + A_4 c_{ggh}^2 c_{hhh}^2 + A_5 c_{gghh}^2 + A_6 c_{tt} c_t^2 + A_7 c_t^3 c_{hhh} \\ & + A_8 c_{tt} c_t c_{hhh} + A_9 c_{tt} c_{ggh} c_{hhh} + A_{10} c_{tt} c_{gghh} + A_{11} c_t^2 c_{ggh} c_{hhh} + A_{12} c_t^2 c_{gghh} \\ & + A_{13} c_t c_{hhh}^2 c_{ggh} + A_{14} c_t c_{hhh} c_{gghh} + A_{15} c_{ggh} c_{hhh} c_{gghh} + A_{16} c_t^3 c_{ggh} + A_{17} c_t c_{tt} c_{ggh} \\ & + A_{18} c_t c_{ggh}^2 c_{hhh} + A_{19} c_t c_{ggh} c_{gghh} + A_{20} c_t^2 c_{ggh}^2 + A_{21} c_{tt} c_{ggh}^2 + A_{22} c_{ggh}^3 c_{hhh} \\ & + A_{23} c_{ggh}^2 c_{gghh}. \end{aligned} \quad (7.2)$$

The coefficients A_i are evaluated in bins of width 20 GeV from 250 GeV to 1050 GeV, i.e. for 40 bins. The median of the statistical uncertainties of the differential coefficients A_i does not exceed 3%, however in the bins beyond $m_{hh} \gtrsim 650$ GeV some A_i coefficients have uncertainties in the 20-30% range.

7.1.1 Definition of shape types

We distinguish four types of characteristic shapes for the Higgs boson invariant mass distribution m_{hh} :

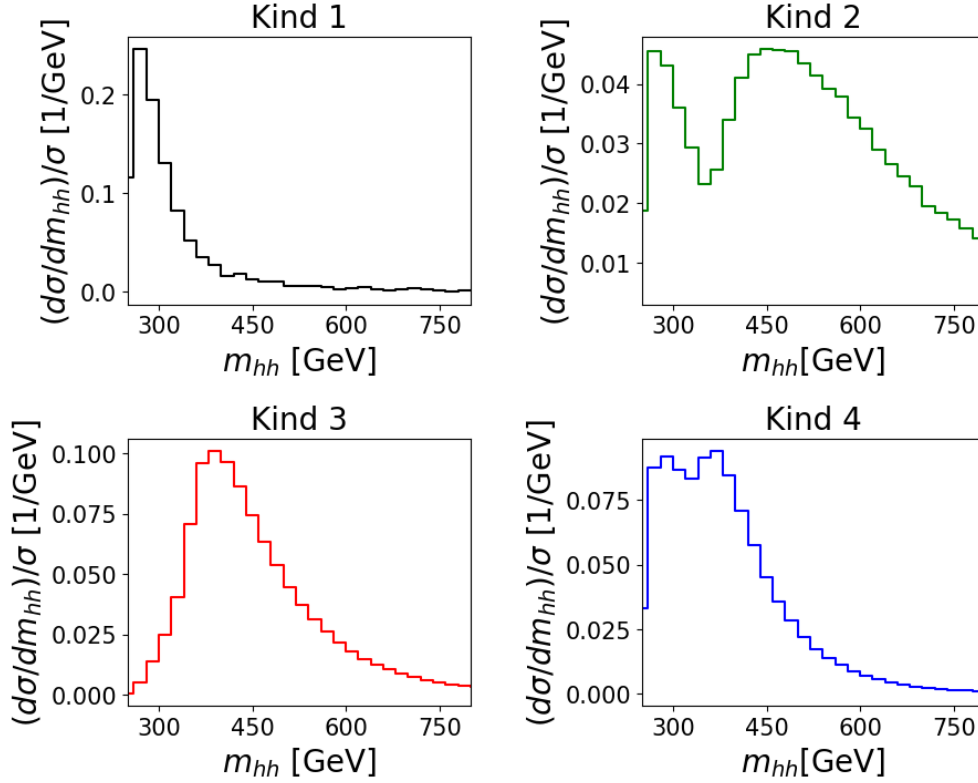


Figure 7.1: The four kinds of shapes defined in our analyzer to classify the m_{hh} distributions. The colours correspond to the colours shown in Figs. 7.2 to 7.6.

1. Enhanced low m_{hh} region, constantly falling distribution as m_{hh} increases.
2. Double peak with peaks separated by more than 100 GeV.
3. Single peak near the $t\bar{t}$ production threshold at $m_{hh} \sim 346$ GeV.
4. Double peak with peaks separated by less than 100 GeV.

Examples of the four shape types are shown in Fig. 7.1. According to our classification the Standard Model shape is contained in distributions of kind 3. Certainly there is some arbitrariness in the definition of these shapes. For example, shapes of kind 4 would move to kind 1 or 3 for bin widths ≥ 100 GeV. However, the other three shape types are quite robust and would be clearly distinguishable experimentally.

Based on the parametrisation in Eq. (7.2), the normalised differential cross section is computed for a 5-dimensional grid in the coupling parameter space and according to its behaviour is classified into one of the four shape types. For this purpose we wrote a function, called “analyzer” in the following, that checks the slopes of the

distribution and puts it into the corresponding class. At this stage the shape classes are mutually exclusive. For each point in the coupling parameter space, we also consider the variations of the result in each bin due to inclusion of the statistical uncertainties on the coefficients A_i . If the shape obtained after these variations belongs to a different kind, we exclude that point from the data set. We find that for shape type 4 about 20% of points fall into this category and are therefore excluded, while for shape type 2 it is about 8%, and for types 1 and 3 it is less than 5%. Scale variations have not been included, as they tend to be rather uniform over the whole m_{hh} range [184] and therefore would not significantly modify our shape analysis.

7.1.2 Classification of m_{hh} distributions

Our results for the $gg \rightarrow hh$ cross sections at NLO are produced for a centre-of-mass energy of $\sqrt{s} = 13$ TeV, using PDF4LHC15_nlo_100_pdfas parton distribution functions interfaced via LHAPDF, along with the corresponding value for α_s . The masses have been set to $m_h = 125$ GeV, $m_t = 173$ GeV and the top quark width has been set to zero.

We study the differential cross section as a function of five anomalous couplings, varying them in the ranges specified below,

$$c_t \in [0.5, 1.5], c_{hhh} \in [-3, 8], c_{tt} \in [-3, 3], c_{ggh}, c_{gggh} \in [-0.5, 0.5]. \quad (7.3)$$

The ranges are motivated by current experimental constraints. For c_{hhh} we use a smaller range than suggested by experiment in order to focus more on the range where interesting shape features are present. In order to visualise the results, we project out 2-dimensional slices of the 5-dimensional parameter space, fixing the other three couplings to their SM values. This leads to a total of ten configurations. For each of these ten projections we generated a set of 10^6 parameter pairs. Feeding them through our analyzer we obtain the shape type produced by the given point in the coupling parameter space. The results are shown in Figs. 7.2–7.6. The white diamonds denote the Standard Model point in the parameter space.

In Fig. 7.2 we display variations of the top quark Yukawa coupling c_t versus the trilinear Higgs coupling c_{hhh} (left) and the effective gluon-Higgs couplings, c_{ggh} versus c_{gggh} (right). In all the figures where two couplings are varied, the other three couplings are set to their SM values. It can be clearly seen that the shapes of kind 1, i.e. shapes with an enhanced low m_{hh} region (marked in black), are resulting from larger c_{hhh} values. The total cross section as a function of c_{hhh} is a parabola with a minimum around $c_{hhh} \approx 2.4$, while for $|c_{hhh}| \gtrsim 3$ and $c_t = 1$ the distribution is enhanced in the low m_{hh} region, where the triangle-type contributions dominate. Larger/smaller values of c_t shift this behaviour towards larger/smaller values of c_{hhh} because they enhance/decrease the box-type contributions. For shapes of kind 2, i.e. double peaks with a separation of more than 100 GeV (green), we find that such a shape can be produced for coupling values which are rather close to the SM values. Shapes of kind 3 (red) are SM-like. They only cover about one quarter of the $c_t - c_{hhh}$ plane. Shapes of kind 4 (blue) have a double peak separated by less than 100 GeV. For $c_{tt} = c_{ggh} =$

7.1 Classification through predefined shape types

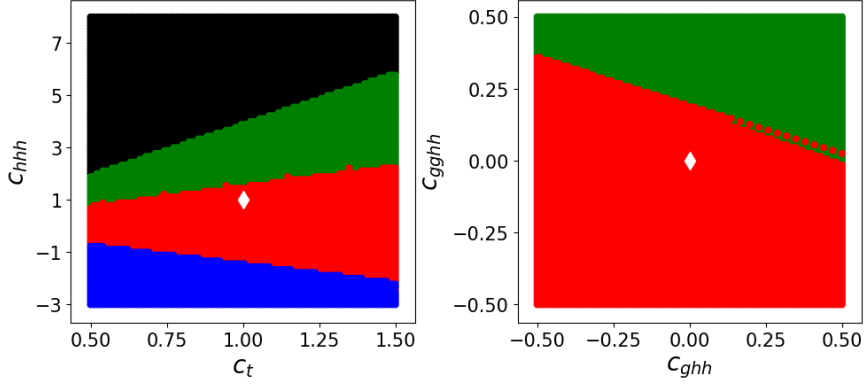


Figure 7.2: The parameter regions leading to each predefined shape type in the $c_t - c_{hhh}$ (left) and $c_{ggh} - c_{gghh}$ (right) parameter spaces. The black area denotes shapes of kind 1 (enhanced low m_{hh} region; green: kind 2 (well separated double peaks), red: kind 3 (SM-like), blue: kind 4 (close-by double peaks). The white diamonds mark the Standard Model point.

$c_{gghh} = 0$, such structures only occur for negative values of c_{hhh} , over the whole allowed c_t range.

Considering variations of c_{ggh} versus c_{gghh} , shown in the right-hand panel of Fig. 7.2, we find only shapes of kind 2 (green) and SM-like shapes (red). The existence of kind 2 shapes means that a double peak structure could be produced solely by effective Higgs-gluon couplings, while keeping c_{hhh}, c_t and c_{tt} at their SM values. However, for the more likely case that c_{ggh} deviates only slightly from zero [213], and so does c_{gghh} , these couplings do not distort the SM shape significantly.

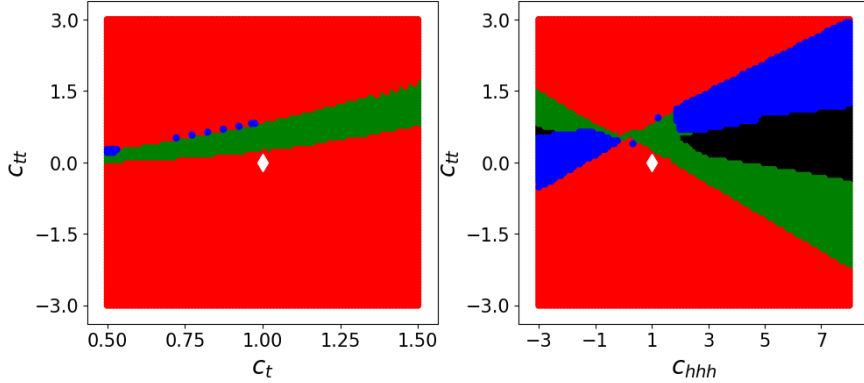


Figure 7.3: The parameter regions associated to each shape type in the $c_t - c_{tt}$ (left) and $c_{hhh} - c_{tt}$ (right) parameter spaces. For the colour code we refer to Fig. 7.1.

Variations of c_t versus c_{tt} and c_{hhh} versus c_{tt} are shown in Fig. 7.3. Varying only c_t and c_{tt} , the shapes remain mainly SM-like. A small area in the $c_t - c_{tt}$ plane however contains doubly peaked m_{hh} distributions, which thus can originate from anomalous

top-Higgs couplings only, while the trilinear Higgs coupling remains fixed at its SM value.

Turning to c_{hhh} versus c_{tt} , displayed in the right-hand panel of Fig. 7.3, we find that for kind 1 and kind 4 shapes the parameter regions are split into two disconnected parts. While shapes of kind 1 are favoured by large values of c_{hhh} , it becomes clear that large values of c_{tt} , also related to triangle-type diagrams, can counterbalance this effect, because the top right corner is not a parameter region producing shapes of kind 1. If both c_{hhh} and c_{tt} are large, it is more likely to produce a double peak structure with close-by peaks (kind 4, blue). Further we see that shapes of kind 2 (well separated double peak structure, green) can be produced by values of c_{tt} and c_{hhh} which are rather close to the SM values.

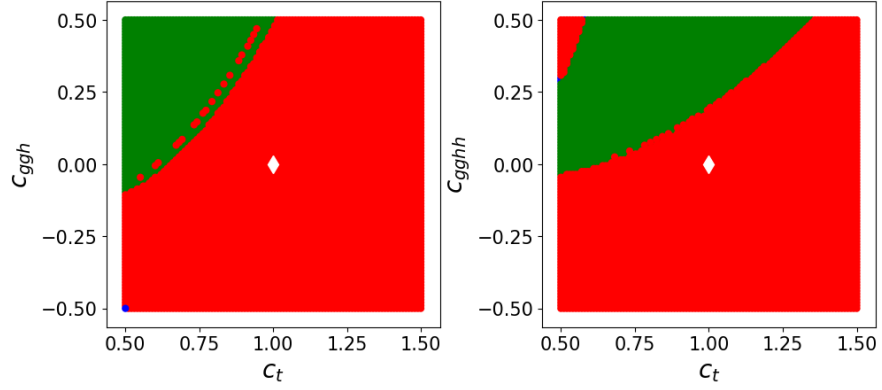


Figure 7.4: The parameter regions associated to each shape type in the $c_t - c_{ggh}$ (left) and $c_t - c_{gghh}$ (right) planes.

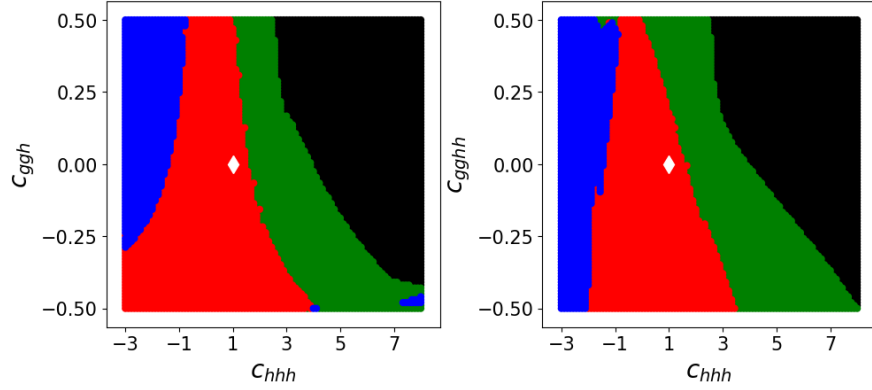


Figure 7.5: The parameter regions associated to each shape type in the $c_{hhh} - c_{ggh}$ (left) and $c_{hhh} - c_{gghh}$ (right) planes.

Fig. 7.4 shows variations of c_t versus c_{ggh} (left) and c_t versus c_{gghh} (right). The parameter space is dominated by SM-like shapes (red), however double peaks can occur

as well (green). We also see that c_{gghh} acts similarly to c_{ggh} in what concerns the shape.

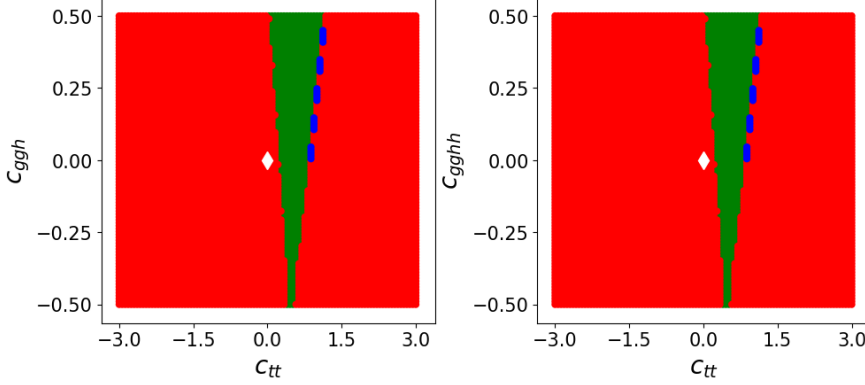


Figure 7.6: The parameter regions associated to each shape type in the $c_{tt} - c_{ggh}$ (left) and $c_{tt} - c_{gghh}$ (right) planes.

For variations of c_{hhh} versus c_{ggh} , shown in Fig. 7.5 (left), all four shape types can occur. The parameter region related to kind 1 (enhanced low m_{hh} , black) is at high values of c_{hhh} as expected, and the kind 2 shapes (well separated double peak, green) can be seen as a transition from kind 1 to kind 3. Close-by double peaks (kind 4, blue) however are mostly associated to negative c_{hhh} values. Note that a similar pattern can be found in Fig. 7.2 (left). Variations of c_{hhh} versus c_{gghh} , shown in Fig. 7.5 (right), are similar in the overall behaviour, and again show that c_{gghh} and c_{ggh} have a similar impact on the shape.

Fig. 7.6 shows variations of c_{tt} versus c_{ggh} (left) and c_{tt} versus c_{gghh} (right). We observe that SM-like shapes (red) are preferred. However, doubly peaked structures are also possible for c_{tt} values not too far from the SM value ($c_{tt} = 0$). We also notice the similarity to Fig. 7.3 (left). The behaviour with respect to c_{gghh} is again similar.

Note that in SMEFT, c_{ggh} and c_{gghh} are related, so this behaviour would necessarily be the case. However we will see later that a shape classification algorithm based on unsupervised learning is able to detect shape differences which distinguish effects of c_{ggh} and c_{gghh} . An interesting feature is also that kind 1 (black) and kind 4 (blue) shapes appear only when we modify the value of c_{hhh} : for $c_{hhh} = 1$ shapes of kind 1 never occur, and shapes of kind 4 are very unlikely. Further, the kind 4 shapes tend to point to (moderately) negative values of c_{hhh} as long as c_{tt} is close to zero, as can be seen from Figs. 7.2, 7.3 and 7.5.

7.1.3 Classification of $p_{T,h}$ distributions

So far we have studied m_{hh} distributions, assuming that they are very well suited to study the sensitivity to shape changes induced by anomalous couplings. In order to verify that we do not miss out interesting features in the transverse momentum distributions, we also present a study of the $p_{T,h}$ distributions, but only at LO, to

assess the salient features. The main difference with respect to the m_{hh} case is that in the $p_{T,h}$ analysis we could identify only two kinds of clearly distinct shapes: single peak (SM-like, which we denote as ‘ $p_{T,h}$ kind 3’) and double peak, with peaks separated by at least 30 GeV (denoted as ‘ $p_{T,h}$ kind 2’). Examples of such $p_{T,h}$ shapes are shown in Fig. 7.7.

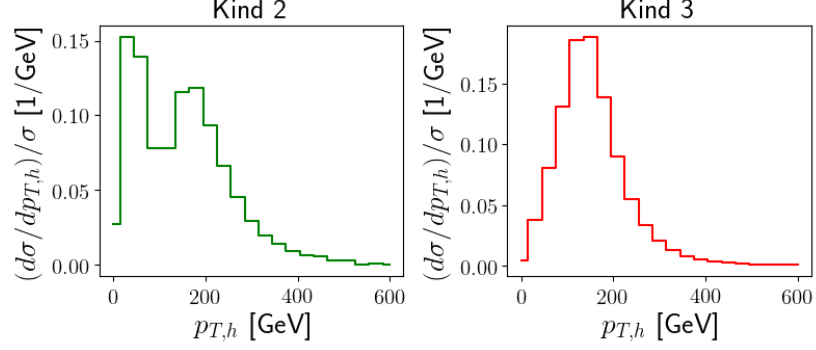


Figure 7.7: Examples of $p_{T,h}$ distributions with a single or double peak.

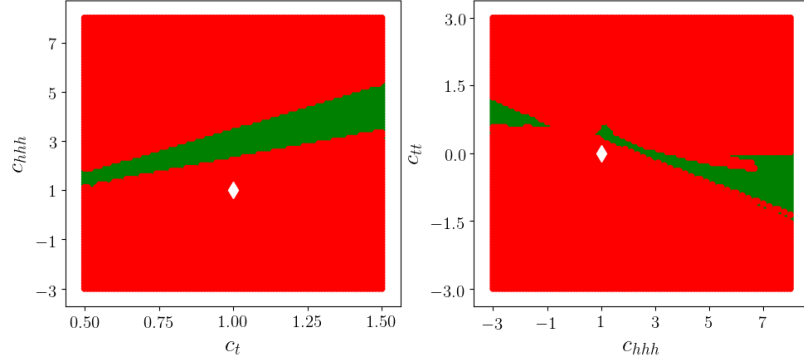


Figure 7.8: The parameter space associated to each shape type in the $c_t - c_{hhh}$ and $c_{hhh} - c_{tt}$ planes for the $p_{T,h}$ distribution.

The parameter spaces leading to singly or doubly peaked shapes are shown in Fig. 7.8 for the $c_t - c_{hhh}$ and $c_{hhh} - c_{tt}$ configurations. The parameter region related to shapes with a well separated double peak (green) is similar to the m_{hh} case, as one can see comparing with Figs. 7.2 and 7.3. This indicates that the underlying parameter space leads to similar characteristics for the distributions differential in $p_{T,h}$ and m_{hh} , however the $p_{T,h}$ distribution is less sensitive than the m_{hh} distribution.

7.2 Classification and clustering by unsupervised learning

7.2.1 Unsupervised learning procedure

To assess the bias introduced by the definition of the four shape types, and to find a more flexible classification which can be extended easily to more than four shape types, we approach the classification problem using unsupervised learning techniques. We construct a classification of the shapes of the m_{hh} distribution into distinct types, where we do not predefine what the types should look like. For this purpose we use a Multilayer autoencoder to find common patterns in the data and thus achieve a compressed latent space representation. The setup is implemented using KERAS [214] and TENSORFLOW [215]. As input data we use 30 bins of width 20 GeV for the normalised m_{hh} distributions. We train the network based on a set of 10^5 distributions, retaining 10% for the validation. The encoder architecture, i.e. the part compressing the array information, is composed of two dense layers with 20 nodes and a middle layer with 4 nodes, the latter defining the length of the array containing the compressed information. The decoder architecture, which reconstructs the original array from the compressed one, is composed of two dense layers of 20 nodes and an output layer of the length of the input array.

To test how stable our results are against variations of the training data set and the encoding procedure, and to reduce uncertainties, for example due to overfitting, we produced ten different autoencoder models. For each model we picked 10^4 random points from the training set for validation to start from different training and validation sets and a different initialization of the weights. We trained the autoencoder for each model over 10000 epochs using Adam [216] as optimizer and the root mean square error to define the loss function. Based on the trained autoencoder we applied the encoder models to the training and validation data to obtain two sets of compressed arrays for each of the ten models. The ten different encoded training data sets are then fed to a classification algorithm, where we employed the KMeans clustering algorithm from scikit-learn [211], asking for a classification into a given number of clusters. We tested classifications into four to eight clusters.

Asking the KMeans algorithm to find four clusters yielded the shape types shown in Fig. 7.9, the result of asking for seven clusters is shown in Fig. 7.10. The curves denote the cluster centres determined by the KMeans algorithm, for each of the ten encoder models, with a colour code as defined in Table 7.1.

One can see from Figs. 7.9 and 7.10 that in the case of clustering into four shape types, cluster 2 contains shapes which vary substantially. In contrast, for seven shape types, the cluster centers obtained from the ten different encoder models are quite similar. Asking for 5–8 clusters we found that seven clusters seemed to be the optimal number to capture distinct shape features, while defining eight clusters did not lead to useful additional features but rather to the tendency to focus on local minima in the clustering space, while neglecting more global shape features.

The four clusters shown in Fig. 7.9 do only partly coincide with the ones defined in Section 7.1.1. Shapes of kind 1, showing an enhanced m_{hh} region, as well as shapes of

7 Higgs boson pair production and m_{hh} distribution shape analysis

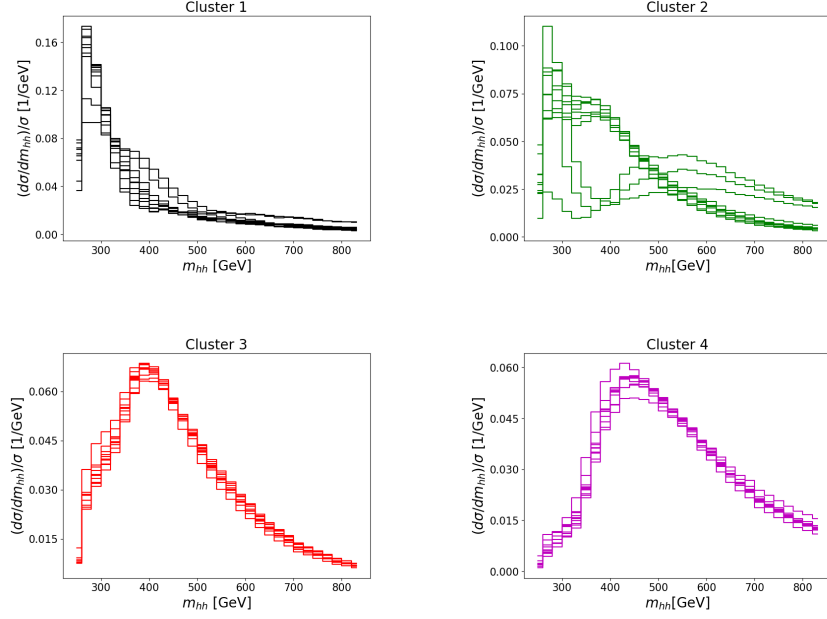


Figure 7.9: The clusters obtained by asking for a classification into four shape types. We show the cluster centres obtained from 10 different encoder models, in the colour code defined in Table 7.1.

kind 3 (SM-like), were clearly identified. Shapes having a double peak were clustered together with shapes showing a shoulder. However, a cluster was formed which was not considered in the predefined types, containing shapes with an enhanced tail.

To combine the results from the ten clustering procedures, we adopted the “majority vote” method, i.e. for each of the ten clustering procedures a given point in the coupling parameter space gets a label (“vote”) corresponding to the cluster it belongs to. The final cluster assigned to that point is the one which collected the largest number of votes.

7.2.2 Parameter space underlying the clusters

In this section we show how the parameter space relates to the clusters if we ask for four or seven clusters. For each parameter configuration of our 5-dimensional grid, we plot the corresponding cluster type in Fig. 7.11 and Figs. 7.13 to 7.16. The colour codes are shown in Figs. 7.9 and 7.10, and are also listed in Table 7.1. For clusters which are similar to the shape types defined in Section 7.1.1, we should also find patterns similar to the ones shown in Figs. 7.2 to 7.6.

Comparing Fig. 7.11 (top row) with Fig. 7.2 (left), both showing variations of c_t versus c_{hh} , we see that kind 1 shapes (black) are clearly identified. However, for

7.2 Classification and clustering by unsupervised learning

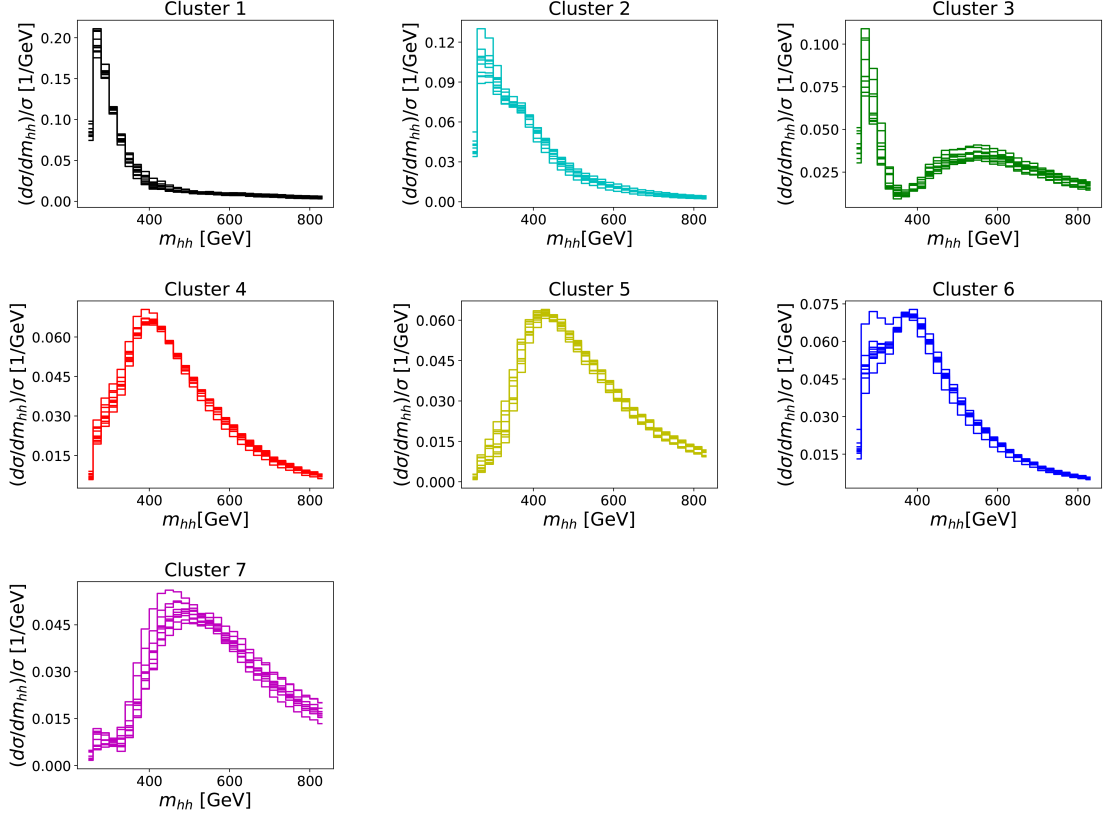


Figure 7.10: The clusters obtained by asking for a classification into seven shape types. The cluster centres obtained from 10 different encoder models are shown in the colour code defined in Table 7.1.

both four and seven clusters the area for SM-like shapes got smaller, as the clustering algorithm also identifies features which were not considered in the predefined shapes. For example, the clustering into four clusters identifies shapes which are almost SM-like but have an enhanced tail (magenta), and the clustering into seven clusters in addition identifies shapes which are almost SM-like but have a shoulder (blue). Certainly we could have defined such features in our analyzer as well, but it is not that easy to define where the tail starts and what exactly should be considered as “enhanced”. Further, the figure clearly shows that small variations of c_{hhh} can easily distort the SM-like shape, while the shape is more robust against variations of c_t . Fig. 7.11 (bottom row) shows c_t versus c_{tt} . We again see that variations c_t and c_{tt} mostly produce SM-like shapes. Why this is so can be understood from the behaviour of the coefficients A_i in eq. (7.2)

Cluster	closest predefined type	colour
4 clusters		
1	kind 1 (enhanced low m_{hh})	black
2	double peak/enhanced tail	magenta
3	kind 3 (SM-like)	red
4	kind 4 (close-by double peaks)/shoulder	blue
7 clusters		
1	enhanced low m_{hh}	black
2	enhanced low m_{hh} , slowly falling or shoulder	cyan
3	enhanced low m_{hh} , second local maximum above $m_{hh} \simeq 2m_t$	green
4	SM-like	red
5	SM-like with enhanced tail	yellow
6	close-by double peaks or shoulder left	blue
7	no steep slope at low m_{hh} , enhanced tail	magenta

Table 7.1: Clusters and shape types with corresponding colour codes for the classification into four and seven clusters.

which are relevant in these cases. For Fig. 7.11 (top row), only the coefficients A_1, A_3 and A_7 are relevant. As A_1 and A_7 have opposite signs and a different peak location, this can generate a rich shape structure. For Fig. 7.11 (bottom row), the coefficients A_2, A_6 and A_8 are relevant in addition to A_1, A_3 and A_7 . A_2 being the coefficient of c_{tt}^2 , it is dominant except for very small values of c_{tt} and leads to a SM-like shape. We also observe that c_{tt} has the tendency to enhance the total cross section, such that only a relatively small slice in c_{tt} is left after considering the bounds on the total cross section.

Fig. 7.13 (top row) shows c_{hhh} versus c_{tt} , where we see that the interplay between c_{hhh} and c_{tt} can lead to all shape types. Comparing Fig. 7.13 (bottom row) with Fig. 7.2 (right), showing variations of c_{ggh} versus c_{gghh} , we observe that the unsupervised learning algorithm with seven clusters distinguishes four shape types, showing that large values of c_{ggh} and c_{gghh} favour shapes with an enhanced tail (magenta) or/and a double peak (green), while negative values favour a shoulder on the left of the peak (blue). The limits on the total cross section do not exclude any parameter range in this panel.

A behaviour similar to the one in Fig. 7.11 can be seen in Fig. 7.14: as c_{hhh} varies the distribution goes through various shape types, while variations of c_{ggh} and c_{gghh} affect the shapes to less extent. Fig. 7.14 also shows that a positive c_{gghh} value has the tendency to enhance the tail of the distribution.

Fig. 7.15 shows c_{tt} versus c_{ggh} (top) and c_{tt} versus c_{gghh} (bottom). Compared to Fig. 7.6, the clustering into both four and seven clusters shows a better discrimination power between SM-like shapes and small deviations, for example due to an enhanced

7.2 Classification and clustering by unsupervised learning

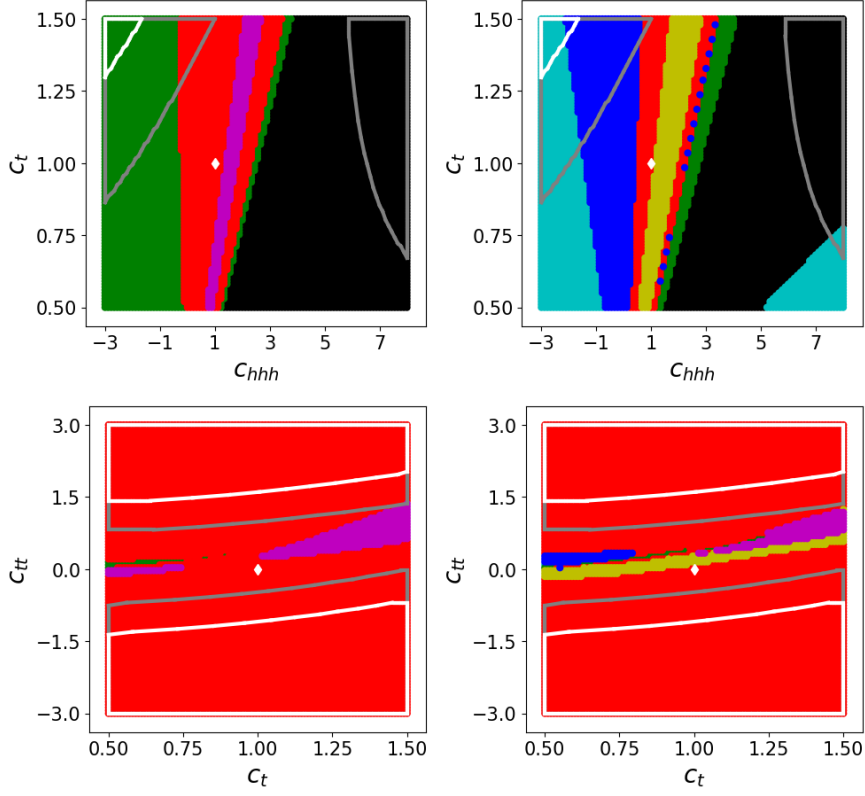


Figure 7.11: Shape types produced by variations of c_t versus c_{hhh} (top) and c_t versus c_{tt} (bottom). Left: 4 clusters, right: 7 clusters. The areas outside the silver and white curves are regions where the total cross section exceeds $6.9 \times \sigma_{SM}$ and $22.2 \times \sigma_{SM}$, respectively. These values are motivated by the current ATLAS/CMS limits at $\sqrt{s} = 13$ TeV [32, 33]. The red areas denote SM-like shapes. The full colour code is given in Table 7.1.

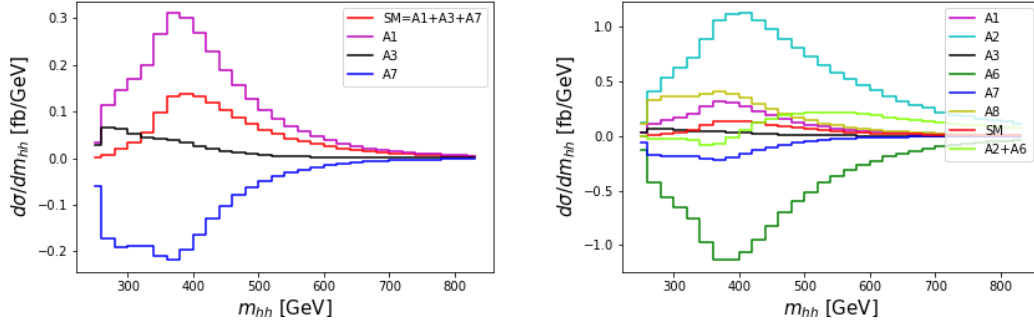


Figure 7.12: Contributions of the coefficients A_i in eq. (7.2) which are relevant for Fig. 7.11.

tail. We again see that c_{tt} has a larger impact on the total cross section than c_{ggh} or c_{gghh} .

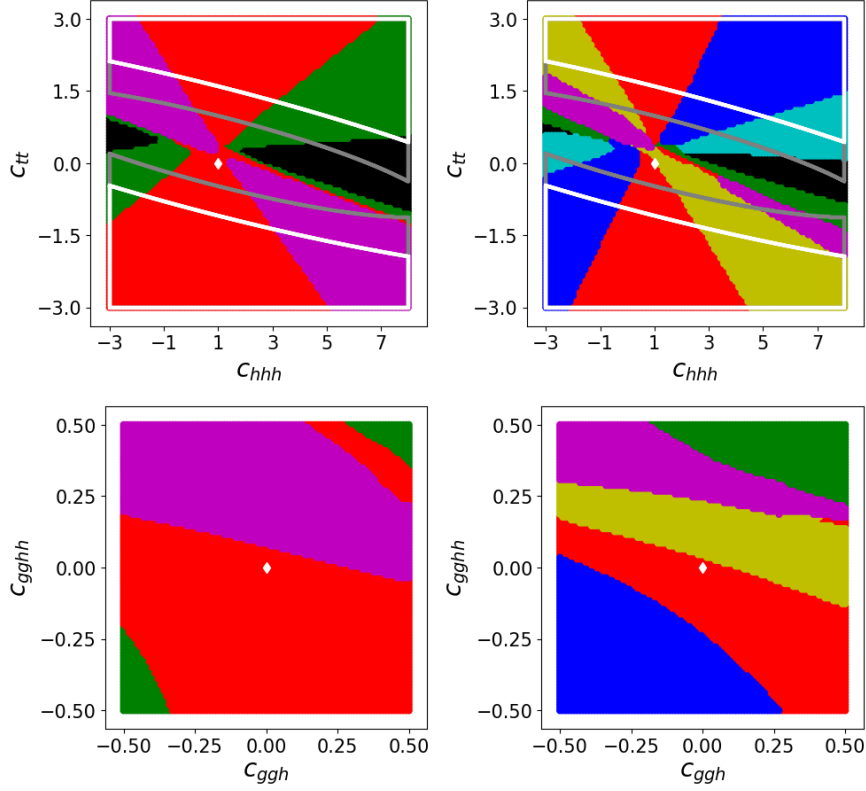


Figure 7.13: Shape types produced by variations of c_{hhh} versus c_{tt} (top) and c_{ggh} versus c_{gghh} (bottom). Left: 4 clusters, right: 7 clusters. The areas outside the silver and white curves are regions where the total cross section exceeds $6.9 \times \sigma_{SM}$ and $22.2 \times \sigma_{SM}$, respectively. These values are motivated by the current ATLAS/CMS limits at $\sqrt{s} = 13$ TeV [32, 33]. The red areas denote SM-like shapes. The full colour code is given in Table 7.1.

Fig. 7.16, showing the $c_t - c_{ggh}$ and $c_t - c_{gghh}$ parameter planes, can be compared to Fig. 7.4. Again, both the case of four and of seven clusters indicates that the unsupervised learning algorithm is able to distinguish better subtle influences on the shape than our method based on humanly classified shapes.

In Fig. 7.17, we compare results of our shape analysis produced with LO and NLO input data. We observe that NLO corrections can change the shape considerably and therefore are important for a shape analysis.

The results above have shown that the parameters c_{hhh} and c_{tt} have the largest influence on the shape. In SMEFT, c_{tt} is suppressed compared to c_t by one order of the large new physics scale [61]. Furthermore, SMEFT imposes the relation $c_{gghh} = c_{ggh}/2$. Using this relation and imposing that c_{tt} amounts to 5% of c_t , we obtain a 3-dimensional parameter space simulating the SMEFT situation, which is visualized in Fig. 7.18.

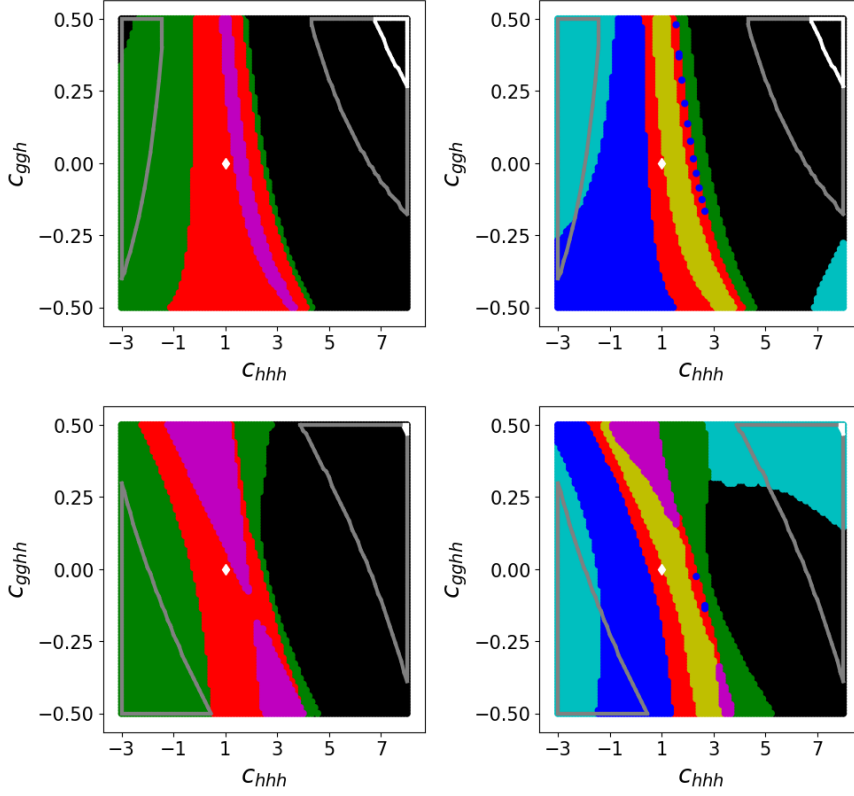


Figure 7.14: Shape types produced by variations of c_{hhh} versus c_{ggh} (top) and c_{gggh} (bottom). Left: 4 clusters, right: 7 clusters. The areas outside the silver and white curves are regions where the total cross section exceeds $6.9 \times \sigma_{SM}$ and $22.2 \times \sigma_{SM}$, respectively. These values are motivated by the current ATLAS/CMS limits at $\sqrt{s} = 13$ TeV [32, 33]. The red areas denote SM-like shapes. The full colour code is given in Table 7.1.

7.3 Conclusions

The aim of this work was to provide more insight how certain configurations of anomalous couplings in the Higgs sector lead to a corresponding characteristic shape of the Higgs boson pair invariant mass distribution. For this purpose we employed the Lagrangian relevant to Higgs boson pair production as given in a non-linear Effective Field Theory framework, which contains five (potentially) anomalous couplings [111]. We produced data for the Higgs boson pair invariant mass distribution, based on a calculation which includes the NLO QCD corrections with full top quark mass dependence, varying all five coupling parameters by finite steps, thus producing a dense grid of data. Then we defined four characteristic shape types for the m_{hh} distribution and visualised the parameter space leading to these shape types. To this aim we projected onto all possible two-dimensional slices of the parameter space, keeping the remaining parameters at their Standard Model values. We also considered $p_{T,h}$ distributions for a

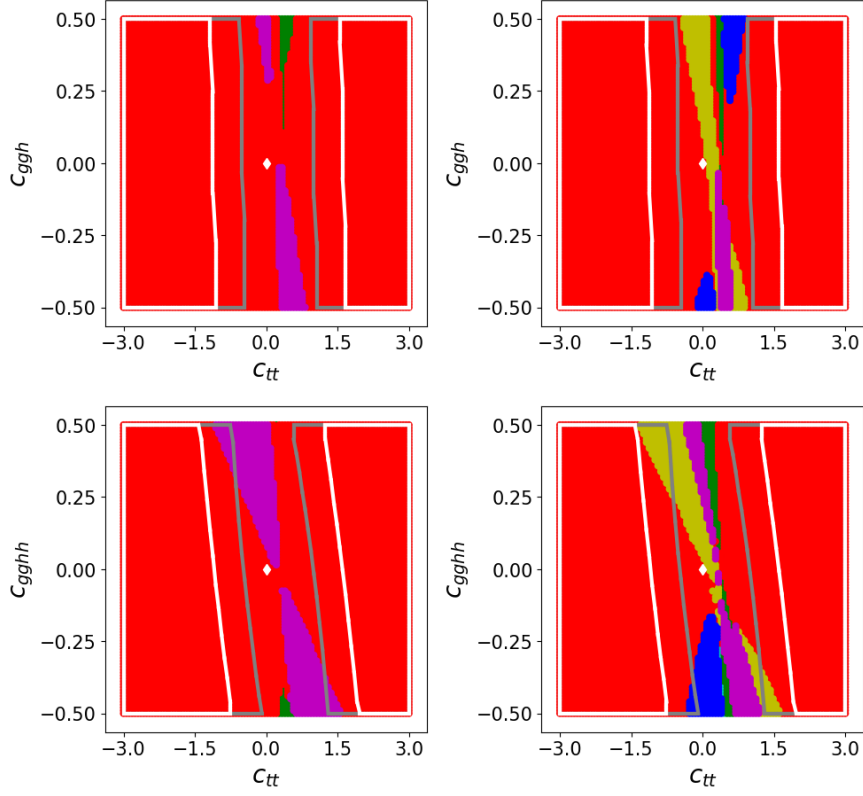


Figure 7.15: Shape types produced by variations of c_{tt} versus c_{ggh} (top) and c_{gghh} (bottom). Left: 4 clusters, right: 7 clusters. The areas outside the silver and white curves are regions where the total cross section exceeds $6.9 \times \sigma_{SM}$ and $22.2 \times \sigma_{SM}$, respectively. These values are motivated by the current ATLAS/CMS limits at $\sqrt{s} = 13 \text{ TeV}$ [32, 33]. The red areas denote SM-like shapes.

shape analysis, however we found that the m_{hh} distribution is more sensitive to shape changes induced by anomalous couplings.

Further, we tested an unsupervised learning approach to classify shapes. We produced 10^5 distributions, trained a neural network based on an autoencoder to extract common shape features and tried to find the number of shape clusters which optimally catches different shape characteristics. Our study demonstrated that some shape features, like an enhanced tail or a shoulder in the m_{hh} distribution, were caught very well by this procedure, and provided more insight about the underlying parameter space leading to such features than the analysis based on predefined shape classes. While machine learning is not essential to define shape clusters, it has the advantage of being easily extendible to a different number of shape types, different binnings or other observables, and of minimising the human bias compared to other shape analysis methods.

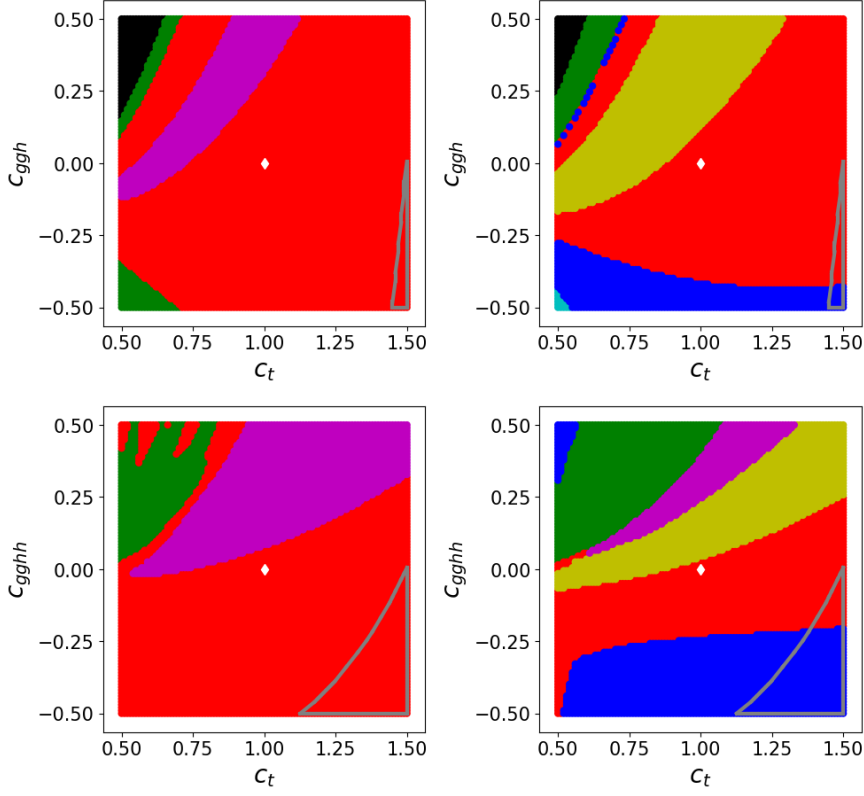


Figure 7.16: Shape types produced by variations of c_t versus c_{ggh} (top) and c_{gggh} (bottom). Left: 4 clusters, right: 7 clusters. The areas outside the silver and white curves are regions where the total cross section exceeds $6.9 \times \sigma_{SM}$ and $22.2 \times \sigma_{SM}$, respectively. These values are motivated by the current ATLAS/CMS limits at $\sqrt{s} = 13$ TeV [32, 33]. The red areas denote SM-like shapes.

The shape analysis revealed that the Standard-Model-like shape is quite stable against variations of c_t , c_{ggh} and c_{gggh} , as long as $c_{hhh} = 1$, while deviations of c_{hhh} from the SM value show a rich shape changing pattern. We also found that small deviations of c_{tt} from zero are very likely to produce a doubly peaked structure in the m_{hh} distribution, while SM-like shapes dominate again as c_{tt} moves further away from zero. However, as c_{tt} leads to a rather fast increase of the total cross section, the shape analysis in combination with the limits on the total cross section allows to put constraints on c_{tt} . This is an interesting feature because, in contrast to c_t and c_{ggh} , c_{tt} cannot be constrained directly from single Higgs boson processes. Further, an enhanced tail or a shoulder of the m_{hh} distribution are likely to be produced by nonzero values of c_{gggh} , however the influence of the effective Higgs-gluon couplings on the shape is milder than the one of c_{hhh} and c_{tt} .

Our approach associates 2-dimensional slices of a 5-dimensional parameter space to a given shape cluster and identifies the directions which are rather flat in parameter

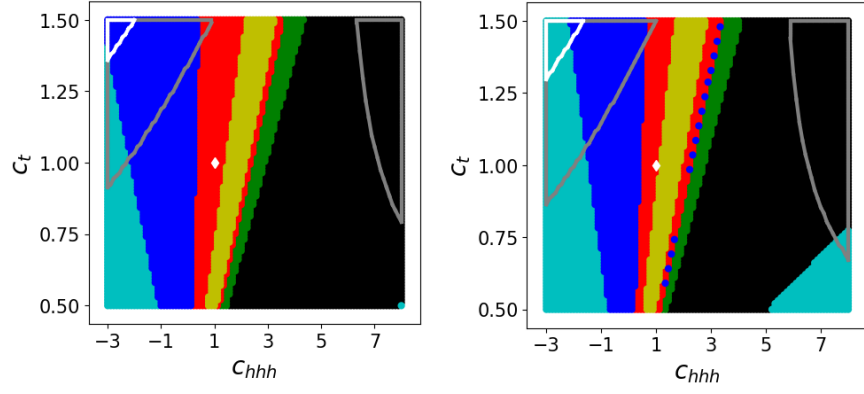


Figure 7.17: Comparison of LO and NLO results for shape types produced by variations of c_t versus c_{hhh} . Left: LO, right: NLO.

space. Therefore it is going beyond a benchmark point analysis, which only provides pointwise snapshots of the 5-dimensional parameter space. The method can also be applied to other processes where anomalous couplings introduce characteristic shape changes to differential cross sections, and it can be extended to consider more than one distribution simultaneously.

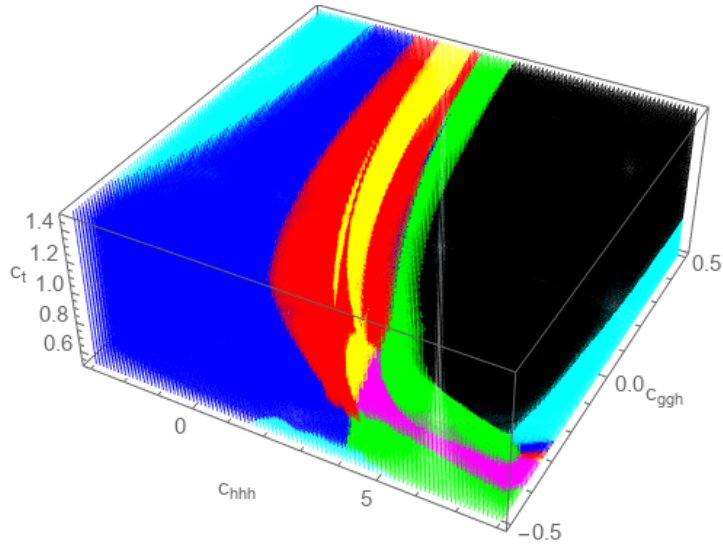


Figure 7.18: Three-dimensional visualisation of shape types produced by variations of c_t , c_{hhh} and c_{ggh} simulating the SMEFT situation. For c_{tt} we used $c_{tt} = 0.05c_t$, while $c_{gggh} = c_{ggh}/2$.

8 Higgs plus one jet production

In this Chapter we present our calculation of the Higgs plus one jet production at NLO QCD with full top mass dependence in the gluon fusion channel. Higgs plus one jet production is an excellent process to study the top Yukawa coupling. As in Chapter 6 the calculation is carried out within the EWChL framework which allows to include EFT effects. The process is characterized by two anomalous couplings c_t and c_{ggh} already presented in Chapter 6. c_t modifies the top Yukawa coupling, responsible of the the top-loop contribution in the gluon fusion channel. The correction includes possible contributions due to top partners. The coupling c_{ggh} instead is parametrizing New Physics heavy colored particles which couple in loops to the gluons and the Higgs boson. The new particles are supposed to be too heavy to be observed at current energy scale. The loop is treated within an effective description resulting in an effective vertex which is weighted by c_{ggh} a coupling that embodies the information about the mass of the new particles.

Direct access to the top Yukawa coupling is given by the $t\bar{t}h$ production, but this process is difficult because its high threshold and the many possible final states. Experimental measurements by the ATLAS and CMS experiments report [5]

$$\sigma_{t\bar{t}h}^{\text{obs}} = 670 \pm 90(\text{stat.})_{-100}^{+110}(\text{syst.}) \quad (8.1)$$

and [30]

$$\frac{\sigma_{t\bar{t}h}^{\text{obs}}}{\sigma_{t\bar{t}h}^{\text{SM}}} = 1.26_{-0.26}^{+0.31} \quad (8.2)$$

respectively, where $\frac{\sigma_{t\bar{t}h}^{\text{obs.}}}{\sigma_{t\bar{t}h}^{\text{SM}}}$ is the observed total cross section normalised to the SM prediction. The experimental readouts are in agreement with the SM but the uncertainties are too big to constrain the top Yukawa coupling properly. More precise measurements will be obtained only after the LHC high luminosity update. The $t\bar{t}h$ experimental status enforces the importance of the Higgs plus one jet production as complementary approach to study the top Yukawa coupling. Even more this process gives an outstanding opportunity to study the nature of the effective Higgs boson coupling to gluons.

Inclusive Higgs production measurements suggest the total cross section of the process to be consistent with the SM prediction [217, 218], in particular the ratio of the total observed Higgs boson signal in gluon fusion yield to the SM predictions is [218]

$$\frac{\sigma_{\text{incl}}^{\text{obs}}}{\sigma_{\text{incl}}^{\text{SM}}} = 1.11_{-0.08}^{+0.09}. \quad (8.3)$$

The total cross section of the process including EFT contributions, according to the Higgs low energy theorem [219] can be approximated as [220]

$$\sigma_{incl}(c_t, c_{ggh}) = (c_t + \frac{3}{2}c_{ggh})^2 \sigma_{incl}^{SM} + \mathcal{O}(\frac{1}{m_t^2}) \quad (8.4)$$

where m_t is top quark mass.

Given the experimental measurements roughly constraining $(c_t + \frac{3}{2}c_{ggh}) = 1$, even if new physics effects are present, they are hard to be observed from the inclusive Higgs production. Modification of the top Yukawa coupling can be compensated by effective ggH contributions and a SM-like inclusive cross section can be obtained even if new physics effects are contributing to the process.

Higgs boson production in association with extra gluon radiation, namely Higgs plus one jet, is a good option to disentangle configuration of the couplings yielding similar total cross sections. It has been shown at LO and NLO HEFT approximation [220, 221] how in the high $p_{T,h}$ region different distributions satisfying $(c_t + \frac{3}{2}c_{ggh}) = 1$ separate and effects due to non SM values of the anomalous couplings c_t and c_{ggh} can be observed. Then the study of the $p_{T,h}$ distribution is an excellent way to observe if there are deviations from the SM due to New Physics contributions. By the way the $p_{T,h}$ value can not be too high, in order to keep valid the effective description of the New Physics.

In this Chapter we present our results where we calculate the total cross section and $p_{T,h}$ distribution for a point on the slice $(c_t + \frac{3}{2}c_{ggh}) = 1$ of the anomalous coupling space, including NLO QCD corrections with full top mass dependence and compare these results to the SM predictions.

Our calculation is implemented within the POWHEG-BOX-V2 [222–224]. The code will be publicly available, allowing the user to choose the values of the anomalous couplings and to match the NLO calculation to a parton shower. The chapter is organized as follows. In Section 8.1 we review the theoretical calculation of Higgs boson production. The Higgs plus one jet production within the EWChL is shown in Section 8.2 including NLO corrections in Section 8.3. Section 8.4 shows phenomenological results. For a point satisfying the condition $c_t + \frac{3}{2}c_{ggh} = 1$, we present the NLO QCD total cross sections normalised to the SM predictions for different values of the cut on $p_{T,h}$ along with its Higgs boson transverse momentum distribution.

8.1 Higgs boson production state of art

Corrections to inclusive Higgs boson production are known up to N³LO QCD accuracy within the HEFT approximation [225]. Full theory corrections are known at LO for H plus one jet [226, 227], Higgs plus 2 jets [228, 229] and Higgs plus three jets [230, 231]. H+1j production fully differential correction within the HEFT approximation are known at NNLO QCD accuracy [232–235], supplemented by the $1/m_t^2$ expansion at NLO QCD accuracy [236, 237] and combined with the exact Born and real corrections [238]. The full theory NLO QCD cross section for H plus 1 jet along with the Higgs

$p_{T,h}$ distribution have been computed in [239] based on a semi-numerical approach. Recently the full set of master integrals relevant for the two-loop NLO QCD virtual corrections have been calculated in [240–243].

Within theories including extra scalars the inclusive Higgs cross section is known to N³LO QCD accuracy in the HEFT approximation [244, 245]. Concerning Higgs plus one jet production, the study of the $p_{T,h}$ spectrum including dimension six and eight operators has been considered in [220, 221, 246–251].

8.2 Non-Linear EFT for Higgs plus one jet production

In the SM model the process is loop-induced, therefore at LO diagrams coming from \mathcal{L}_2 and \mathcal{L}_4 , the LO and NLO Lagrangians in the chiral expansion, have to be taken in account. The relevant terms from the effective Lagrangian $\mathcal{L}_2 + \mathcal{L}_4$ (see Sec. 4.2) are given by:

$$\mathcal{L} \supset -m_t c_t \frac{h}{v} \bar{t}t + \frac{\alpha_s}{8\pi} c_{ggh} \frac{h}{v} G_{\mu\nu}^a G_a^{\mu\nu} \quad (8.5)$$

The coupling c_t is from \mathcal{L}_2 , the Higgs-gluon couplings c_{ggh} from \mathcal{L}_4 and are the same of

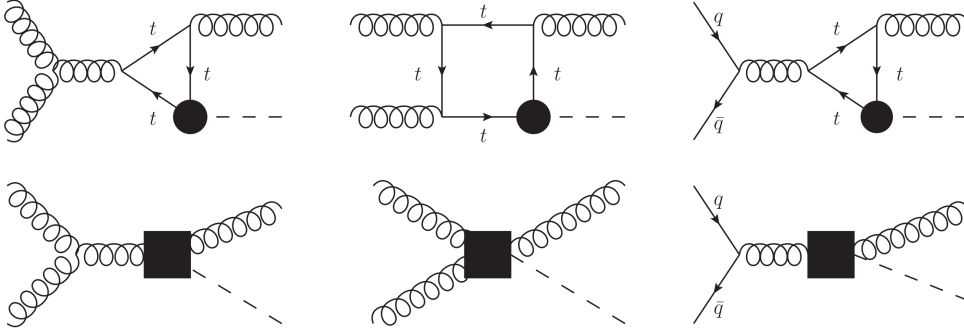


Figure 8.1: Higgs plus one gluon production at leading order in the chiral Lagrangian. The black dots indicate vertices from \mathcal{L}_2 , the black squares denote local terms from \mathcal{L}_4 . We distinguish the light quark and the top quark lines specified by q and t respectively.

Sec. 6.2. We neglect CP-violating effects. The SM is obtained setting $c_t = 1$, $c_{ggh} = 0$. As for Higgs boson pair production all the diagrams are at the same order in the chiral counting (chiral dimension 4, equivalent to one-loop order). The process is mediated by gg , gq and $\bar{q}q$ initial states. Some of the LO diagrams are shown in Fig. 8.1. Defining $\kappa_g = \frac{3}{2}c_{ggh}$ according to the notation of [220] the squared amplitude of the process reads:

$$|\mathcal{M}(c_t, \kappa_g)|^2 = |c_t M_{loop} + \kappa_g M_{tree}|^2 \quad (8.6)$$

8 Higgs plus one jet production

where M_{loop} is the loop-induced SM-like amplitude and M_{tree} the effective amplitude. The total cross section of the process normalized to the SM can be written as

$$R^{p_{T,h}^{cut}}(c_t, \kappa_g) = \frac{\sigma_{p_{T,h}^{cut}}(c_t, \kappa_g)}{\sigma_{p_{T,h}^{cut}}^{SM}} = [(c_t + \kappa_g)^2 + \delta c_t \kappa_g + \epsilon \kappa_g^2] = [(c_t + \kappa_g)^2 + \Delta(c_t, \kappa_g)] \quad (8.7)$$

where

$$\delta = \frac{2 \int_{p_{T,h}^{cut}}^{\infty} dp_{T,h} d\Omega \text{Re}(\mathcal{M}_{loop} \mathcal{M}_{tree}^*)}{\int_{p_{T,h}^{cut}}^{\infty} dp_{T,h} d\Omega |\mathcal{M}_{loop}|^2} - 2, \quad (8.8)$$

$$\epsilon = \frac{\int_{p_{T,h}^{cut}}^{\infty} dp_{T,h} d\Omega |\mathcal{M}_{tree}|^2}{\int_{p_{T,h}^{cut}}^{\infty} dp_{T,h} d\Omega |\mathcal{M}_{loop}|^2} - 1 \quad (8.9)$$

$$\Delta(c_t, \kappa_g) = \delta c_t \kappa_g + \epsilon \kappa_g^2 \quad (8.10)$$

and $p_{T,h}^{cut}$ is the cut on the Higgs boson transverse momentum, i.e. the minimum value required for the Higgs boson transverse momentum $p_{T,h}$. At LO δ and ϵ have been shown [220] to be very small at low $p_{T,h}$ thus $R_{(c_t+\kappa_g)=1}^{p_{T,h}^{cut}} \sim 1$ for the points belonging to the slice $(c_t + \kappa_g) = (c_t + \frac{3}{2}c_{ggh}) = 1$. δ and ϵ increase with $p_{T,h}$ to become of $\mathcal{O}(1)$ for $p_{T,h}^{cut} > 300$ GeV. This implies that measuring the transverse momentum distribution of the Higgs boson allows to observe effects due to New Physics in contrast to the analysis of the inclusive cross section.

8.3 NLO QCD corrections

We calculate the Higgs plus one jet production in non-linear EFT restricting the NLO corrections to QCD effects. Working within the EWChL framework we include the chiral dimension 6 corrections with QCD coupling g_s^3 . As for the Higgs boson pair production, this procedure is consistent without introducing further anomalous couplings because this effective Lagrangian is renormalisable with respect to QCD.

The leading order amplitude scales as $\sim g_s^3$ therefore the virtual corrections relevant to us are all the diagrams of chiral dimension 6 (loop order two) carrying a factor g_s^5 . They exist as two-loop and one-loop.

The real radiation diagrams proportional to g_s^4 have to be taken in account. As at LO, they can be tree level or one-loop diagrams. In the following we describe the computational setup we used to carry out the calculation.

8.3.1 Leading order

The leading order amplitude has been implemented analytically, the loop-induced SM-like part was adapted from [227] and rescaled by the anomalous coupling c_t , while we computed the tree level EFT part analytically using FORM[252]. As cross check we generated the Born amplitude with GOSAM using the UFO model described in [111]

finding agreement between the two implementation at the amplitude and cross section level.

8.3.2 Virtual corrections

The virtual amplitude of order g_s^5 is composed of the interference of the Born diagrams with two-loop and one-loop diagrams.

The one-loop virtual amplitude has been implemented using the interface between GoSAM and the POWHEG-BOX-V2. At runtime the scalar one-loop integrals involved in the amplitudes were computed using the programs Ninja[253], golem95[254, 255] and OneLoop [256] for the evaluation of the scalar one-loop integrals. For each initial and final state configuration two amplitudes are returned by GoSAM, the first representing the interference of Born tree and virtual one-loop diagrams and the second specifying the interference of Born one-loop and virtual one-loop diagrams. Moreover in the latter case, because of the GoSAM structure we inserted by hand the renormalisation terms, while the in the first case the diagrams are renormalised by the software.

The two-loop amplitude can be decomposed in terms of tensor structures. The $g(p_1)g(p_2) \rightarrow g(p_3)H(p_4)$ amplitude after imposing parity conservation, transversality of the gluon polarization vectors and the Ward identity can be written as the combination of four form factors F_{ijk} multiplying the tensor structures $T_{ijk}^{\mu\nu\tau}$ [257]:

$$\mathcal{M}^{\mu\nu\tau} = F_{212}T_{212}^{\mu\nu\tau} + F_{332}T_{332}^{\mu\nu\tau} + F_{311}T_{311}^{\mu\nu\tau} + F_{312}T_{312}^{\mu\nu\tau} \quad (8.11)$$

where:

$$\begin{aligned} T_{212}^{\mu\nu\tau} &= (s_{12}g^{\mu\nu} - 2p_2^\mu p_1^\nu)(s_{23}p_1^\tau - s_{13}p_2^\tau)/(2s_{13}) \\ T_{332}^{\mu\nu\tau} &= (s_{23}g^{\nu\tau} - 2p_3^\nu p_2^\tau)(s_{13}p_2^\mu - s_{12}p_3^\mu)/(2s_{12}) \\ T_{311}^{\mu\nu\tau} &= (s_{13}g^{\tau\mu} - 2p_1^\tau p_3^\mu)(s_{12}p_3^\nu - s_{23}p_1^\nu)/(2s_{23}) \\ T_{312}^{\mu\nu\tau} &= (g^{\mu\nu}(s_{23}p_1^\tau - s_{13}p_2^\tau) + g^{\nu\tau}(s_{23}p_2^\mu - s_{12}p_3^\nu) + g^{\tau\mu}(s_{12}p_3^\nu - s_{23}p_1^\nu) \\ &\quad + 2p_3^\mu p_1^\nu p_2^\tau - 2p_2^\mu p_3^\nu p_1^\tau)/2 \end{aligned} \quad (8.12)$$

with $s_{ij} = (p_i + p_j)^2$. Three of the form factors are related by cyclic permutations of the external gluon momenta while the fourth is invariant under such permutations. The $q(p_1)\bar{q}(p_2) \rightarrow g(p_3)H(p_4)$ amplitude similarly can be decomposed in terms of two tensor structures as follows [258]

$$M_\rho \epsilon^\rho = F_1 T_1 + F_2 T_2 \quad (8.13)$$

where

$$\begin{aligned} T_1 &= (\bar{u}(p_1)\not{p}_3 v(p_2)p_2 \cdot \epsilon_3 - \bar{u}(p_1)\not{\epsilon}_3 v(p_2)p_2 \cdot p_3), \\ T_2 &= (\bar{u}(p_1)\not{p}_3 v(p_2)p_1 \cdot \epsilon_3 - \bar{u}(p_1)\not{\epsilon}_3 v(p_2)p_1 \cdot p_3) \end{aligned} \quad (8.14)$$

In this case the form factors are related by interchanging the external quark and anti-quark momenta. The form factors can be extracted introducing projectors $P_{\mu\nu\tau}^{ijk}$. The

four projectors for the $gg \rightarrow gH$ amplitude in the D-dimensional space-time are:

$$\begin{aligned}
P_{212}^{\mu\nu\tau} &= \frac{1}{(D-3)s_{23}} \left(-\frac{Ds_{13}}{s_{12}^2 s_{23}} T_{212}^{\mu\nu\tau} + \frac{D-4}{s_{23}^2} T_{332}^{\mu\nu\tau} + \frac{D-4}{s_{12}s_{13}} T_{311}^{\mu\nu\tau} + \frac{D-2}{s_{12}s_{23}} T_{312}^{\mu\nu\tau} \right) \\
P_{332}^{\mu\nu\tau} &= \frac{1}{(D-3)s_{12}} \left(\frac{D-4}{s_{12}s_{23}} T_{212}^{\mu\nu\tau} - \frac{Ds_{12}}{s_{13}s_{23}^2} T_{332}^{\mu\nu\tau} + \frac{D-4}{s_{13}^2} T_{311}^{\mu\nu\tau} + \frac{D-2}{s_{13}s_{23}} T_{312}^{\mu\nu\tau} \right) \\
P_{311}^{\mu\nu\tau} &= \frac{1}{(D-3)s_{13}} \left(\frac{D-4}{s_{12}^2} T_{212}^{\mu\nu\tau} + \frac{D-4}{s_{13}s_{23}} T_{332}^{\mu\nu\tau} - \frac{Ds_{23}}{s_{12}s_{13}^2} T_{311}^{\mu\nu\tau} + \frac{D-2}{s_{12}s_{13}} T_{312}^{\mu\nu\tau} \right) \\
P_{312}^{\mu\nu\tau} &= \frac{D-2}{(D-3)s_{12}s_{13}s_{23}} \left(\frac{s_{13}}{s_{12}} T_{212}^{\mu\nu\tau} + \frac{s_{12}}{s_{23}} T_{332}^{\mu\nu\tau} + \frac{s_{23}}{s_{13}} T_{311}^{\mu\nu\tau} + \frac{D}{D-2} T_{312}^{\mu\nu\tau} \right) \quad (8.15)
\end{aligned}$$

where $P_{\mu\nu\tau}^{ijk} M^{\mu\nu\tau} = F_{ijk}$. For the $q\bar{q} \rightarrow gH$ amplitude the projectors are:

$$\begin{aligned}
P_1 &= \frac{D-2}{2(D-3)s_{12}s_{13}^2} T_1^\dagger - \frac{D-4}{2(D-3)s_{12}s_{13}s_{23}} T_2^\dagger \\
P_2 &= -\frac{D-4}{2(D-3)s_{12}s_{13}s_{23}} T_1^\dagger + \frac{D-2}{2(D-3)s_{12}s_{23}^2} T_2^\dagger \quad (8.16)
\end{aligned}$$

where P_1 and P_2 satisfy $\sum_{spins} P_i M_\rho \epsilon^\rho = F_i$. Contracting the projectors with the amplitude and omitting external spinors, allows to write the amplitude in terms of scalar integrals. The six SM NLO QCD form factors have been computed in [239]. The integrals involved in the form factors have been calculated numerically for a number of phase space points setting $m_H = 125$ GeV and $m_t = 173.055$ GeV. The IR soft and collinear poles of the resulting amplitude have been subtracted to obtain the finite part of the virtual amplitude required in the POWHEG-BOX-V2.

The basis of our calculation are the form factors of [239]. We rescaled the SM form factors by the anomalous coupling c_t . Then using the projectors of eq.s (8.15) and (8.16) we calculated the Born form factors arising from the tree level EFT diagrams and added them to the SM-like Born form factors. This setup gives the virtual two-loop amplitude for a set of precomputed phase space points. In order to call this contribution directly in the POWHEG-BOX-V2 for the 4 channels we created a grid where we stored

$$\beta = \frac{s - m_H^2}{s + m_H^2}, \quad \cos \theta = \frac{t - u}{s - m_H^2}, \quad \text{Amplitude}(\beta, \cos \theta) \quad (8.17)$$

for all the precomputed points, obtaining 4 grids of roughly 2000 points. To interpolate the amplitudes for each channel we used a combination of 10 neural networks trained on bootstrap samples, i.e. subsamples of the initial dataset which can be considered independent and representative of the original dataset. The architecture of each network is composed by three dense layers of nodes 100, 10 and 10 respectively. The setup is implemented using KERAS to produce the models and a modified version of *Keras2cpp* [259] to save them such that they can be loaded in C++. We built a C++ function where, for each channel the 10 models are loaded and the outputs are averaged. This function is called in the POWHEG-BOX-V2 to obtain the two-loop virtual amplitude of any given phase space point. The two-loop virtual amplitude combined with the

one-loop virtual one gives the total virtual amplitude for the process within the EFT framework.

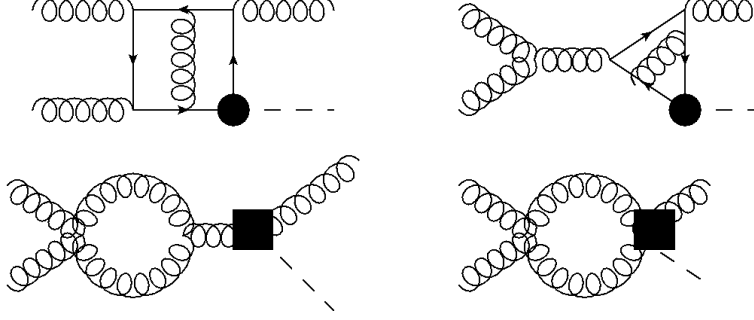


Figure 8.2: Higgs plus one jet production at NLO: Examples for virtual one-loop and two-loop diagrams in the gluon fusion channel.

8.3.3 Real radiation

The real radiation consists of 5-point one-loop topologies with closed top quark loops and tree-level diagrams. Examples of diagrams are shown in Fig. 8.3. The real amplitude squared can be written as

$$|M_{REAL}|^2 = |M_{\text{one-loop}}^{5\text{-point}}|^2 + |M_{\text{tree}}^{5\text{-point}}|^2 + M_{\text{interference}}^{5\text{-point}}. \quad (8.18)$$

The two contributions $|M_{\text{one-loop}}^{5\text{-point}}|^2$ and $(|M_{\text{tree}}^{5\text{-point}}|^2 + M_{\text{interference}}^{5\text{-point}})$ have been generated separately with GoSAM and combined in the POWHEG-BOX-V2. For $(|M_{\text{tree}}^{5\text{-point}}|^2 + M_{\text{interference}}^{5\text{-point}})$ we modified the rescue system to allow the software to rescue amplitudes which are numerically unstable recomputing them in quadruple precision. The amplitudes are computed at run time with Ninja using the quadninja feature for the rescue system. The subtraction of the IR divergences is performed using the FKS method [260].

8.4 Phenomenological results

In this section we present preliminary results produced with the setup described in the previous sections. We produced the NLO QCD total cross section and $p_{T,h}$ differential distribution for the point $(c_t = 0.9, c_{ggh} = \frac{2}{30})$ or equivalently $(c_t = 0.9, \kappa_g = 0.1)$ lying on the slice $(c_t + \frac{3}{2}c_{ggh}) = (c_t + \kappa_g) = 1$ which leads to a degeneracy of the inclusive cross section in the low energy approximation according to equation (8.4).

The top and Higgs masses are set to $m_t = 173.055$ GeV and $m_H = 125$ GeV respectively, to be consistent with the values of the two-loop virtual interpolation, and both the widths are set to zero. The results were computed using PDF4LHC_nlo_30_pdfas

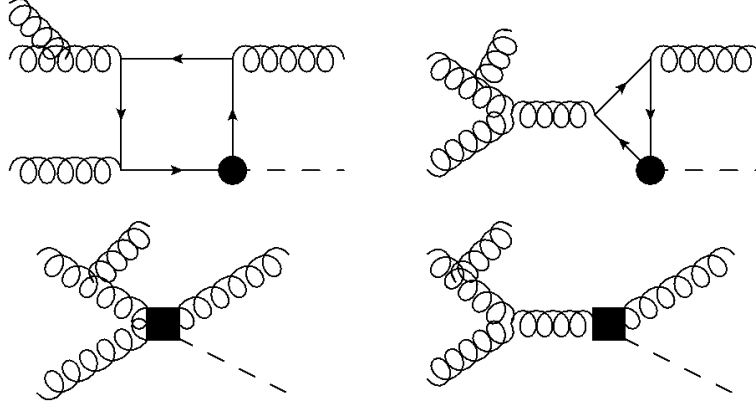


Figure 8.3: Higgs plus one jet production at NLO: Examples for real tree-level and one-loop diagrams in the gluon fusion channel.

	LO [pb]	NLO [pb]	K-factor
$c_t = 0.9, c_{ggh} = \frac{2}{30}$	8.527 ± 0.004	15.690 ± 0.129	1.84 ± 0.01
$c_t = 1, c_{ggh} = 0$ (SM)	8.566 ± 0.001	15.989 ± 0.221	1.87 ± 0.02

Table 8.1: LO and NLO total cross section with statistical error for $c_t = 0.9$, $c_{ggh} = \frac{2}{30}$ and the SM for $p_{T,h} > 30$ GeV. The upper and lower values due to scale variation are shown.

set interfaced via LHAPDF. Jets were clustered using the `anti-kt` jet algorithm implemented in FastJet [261–263] with a radial distance of $R = 0.4$. We used as scales:

$$\mu_F = \mu_R = \frac{H_T}{2} = \frac{1}{2} \left(\sqrt{m_H^2 + p_{T,h}^2} + \sum_i p_{T,i} \right) \quad (8.19)$$

where the sum runs over all the final state partons. All the quarks but the top are treated as massless.

The total cross sections for $(c_t = 0.9, c_{ggh} = \frac{2}{30})$ and the SM at LO and NLO QCD are reported in Table 8.1 along with the K-factor, i.e. the ratio NLO/LO. The table displays how the total cross sections are close, thus also at higher order of the perturbative expansion the $p_{T,h}$ analysis is needed to disentangle possible new physics effects from the SM predictions. As a reflection of the values of the total cross sections also the K-factors are very similar.

The differential transverse momentum distributions of the Higgs boson, for the SM ($c_t = 1, c_{ggh} = 0$) and the BSM configuration ($c_t = 0.9, c_{ggh} = \frac{2}{30}$) are shown in Fig. 8.4.

The red and black curves represent $(c_t = 0.9, c_{ggh} = \frac{2}{30})$ and $(c_t = 1, c_{ggh} = 0)$ respectively where the solid lines are for the NLO QCD distributions while the dotted lines are for the LO ones. At both LO and NLO QCD, in the low $p_{T,H}$ region the two differential cross sections are very close with the BSM curve slightly below the SM one,

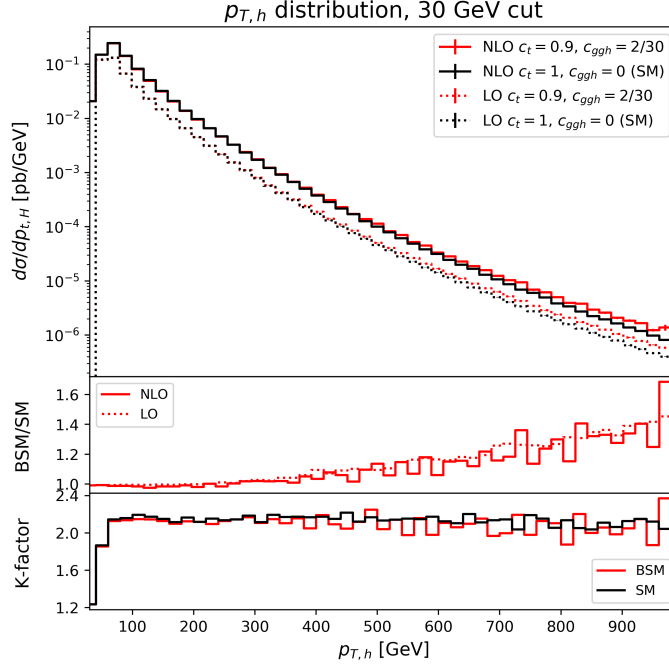


Figure 8.4: Higgs boson transverse momentum spectrum at LO and NLO in QCD. The upper panel shows the differential cross sections, in the middle panel we normalize the BSM distributions to the SM predictions and in the lower panel we show the K-factors (the ratio NLO/LO) for both of the distributions.

they then overlap, they switch around 300 GeV and at high $p_{T,h}$ the BSM distribution is clearly above the SM prediction and the two distributions are disentangled. At LO this trend has been shown already in Refs. [220, 221] and our results are in agreement with them. Even more the middle panel of Fig. 8.4 shows how the ratio BSM/SM is very similar at LO and NLO. The differential K-factors for the two differential cross sections are displayed in the bottom panel of Fig. 8.4. From Ref. [239] it was already known that the SM K-factor is approximately flat for $p_{T,h} \gtrsim 100$ GeV. This peculiar characteristic holds on also including effective deviations from the SM and as already showed at total cross section level, the BSM K-factor is almost degenerate with the SM one. The two K-factors show the large enhancements due to full top mass dependent NLO QCD corrections and how important they are in order to make reliable predictions.

Recalling

$$R_{p_{T,h}}^{p_{T,h}^{cut}}(c_t, c_{ggh}) = \frac{\sigma_{p_{T,h}}^{p_{T,h}^{cut}}(c_t, c_{ggh})}{\sigma_{p_{T,h}}^{SM}} = \left[\left(c_t + \frac{3}{2} c_{ggh} \right)^2 + \Delta(c_t, c_{ggh}) \right] \quad (8.20)$$

8 Higgs plus one jet production

$p_{T,h}^{cut}$ [GeV]	$R^{p_{T,h}^{cut},LO}(0.9, \frac{2}{30})$	$R^{p_{T,h}^{cut},NLO}(0.9, \frac{2}{30})$
30	0.9954 ± 0.0005	0.9837 ± 0.0109
100	1.0004 ± 0.0005	0.9839 ± 0.0013
200	1.0208 ± 0.0011	1.0080 ± 0.0027
300	1.0616 ± 0.0023	1.0419 ± 0.0054
400	1.1131 ± 0.0041	1.0822 ± 0.0093
500	1.1710 ± 0.0048	1.1339 ± 0.0109

Table 8.2: LO and NLO total cross section for $(c_t = 0.9, c_{ggh} = \frac{2}{30})$ normalised to the SM value for different $p_{T,h}^{cut}$, including the statistical error.

the value of the total cross section normalized to the SM prediction, for a point lying on the slice $(c_t + \frac{3}{2}c_{ggh}) = 1$ it reduces to

$$R_{(c_t + \frac{3}{2}c_{ggh})=1}^{p_{T,h}^{cut}}(c_t, c_{ggh}) = [1 + \Delta(c_t, c_{ggh})]. \quad (8.21)$$

We present in Table 8.2, for different cuts on the Higgs boson transverse momentum, $R(0.9, \frac{2}{30})$ at LO and NLO QCD.

Once again we find at LO agreement with [221]. Our analysis displays that, also at NLO QCD, for $p_{T,h}^{cut} = 30 \text{ GeV}$ within the error, the ratio of the total cross sections is compatible with 1, i.e. the total cross section of the BSM configuration is the same as the SM prediction. Increasing $p_{T,h}$ the ratios are moving away from 1. Even though the behaviour of $R^{p_{T,h}^{cut}}(0.9, \frac{2}{30})$ is the same at LO and NLO, $R^{p_{T,h}^{cut},NLO}(0.9, \frac{2}{30})$ is always smaller than $R^{p_{T,h}^{cut},LO}(0.9, \frac{2}{30})$ for each choice of the cut. Focusing on $R^{p_{T,h}^{cut},NLO}(0.9, \frac{2}{30})$ when $p_{T,h}^{cut} \geq 300 \text{ GeV}$ the BSM and SM cross sections can be distinguished. Our NLO QCD results confirmed the importance of the high $p_{T,h}$ analysis in order to observe possible New Physics effects, already known from the LO.

We also present the comparison of NLO QCD full top mass dependent and NLO HEFT $p_{T,h}$ distribution for the point $(c_t = 0.9, c_{ggh} = \frac{2}{30})$. According to eq. (8.4) for all the points on the slice $(c_t + \frac{3}{2}c_{ggh}) = 1$ the HEFT approximation does not distinguish among different configurations of the couplings. The low energy expansion including terms proportional to $1/m_t^2$ reads [219]

$$\sigma(c_t, c_{ggh}) = (c_t + \frac{3}{2}c_{ggh})^2 \left(1 - \frac{7}{15} \frac{\frac{3}{2}c_{ggh}}{c_t + \frac{3}{2}c_{ggh}} \frac{m_h^2}{4m_t^2} \right) \sigma_{SM} + \mathcal{O}(\frac{1}{m_t^4}) \quad (8.22)$$

which allows to take in account deviations due to the anomalous couplings. For a point on the slice $(c_t + \frac{3}{2}c_{ggh}) = 1$ eq. (8.22) reduces to

$$\sigma_{(c_t + \frac{3}{2}c_{ggh})=1}(c_t, c_{ggh}) = (c_t + \frac{3}{2}c_{ggh})^2 \left(1 - \frac{7}{15} \frac{3}{2} c_{ggh} \frac{m_h^2}{4m_t^2} \right) \sigma_{SM} + \mathcal{O}(\frac{1}{m_t^4}). \quad (8.23)$$

In Fig. 8.5 we show our results.

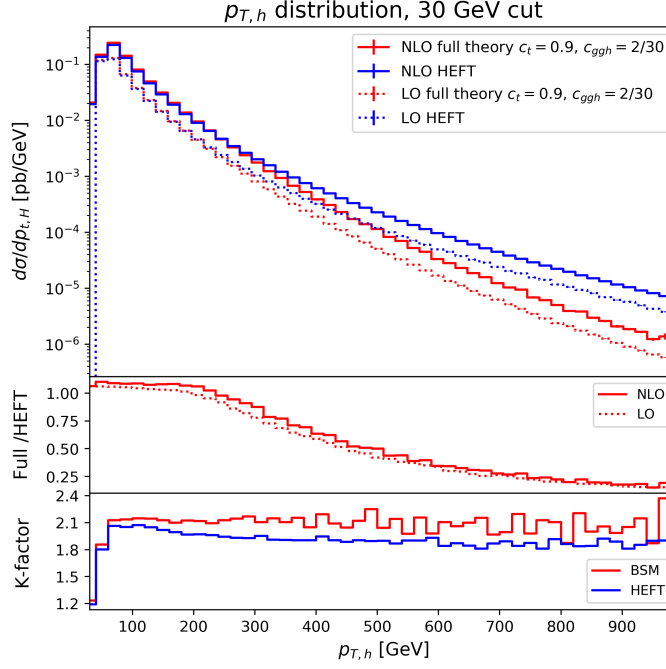


Figure 8.5: Higgs boson transverse momentum spectrum at NLO in QCD with full top mass dependence and in the HEFT approximation. The upper panel shows the differential cross sections, in the middle panel we normalize the BSM distributions to the HEFT predictions and in the lower panel we show the K-factors (the ratio NLO/LO) for both the distributions.

The red and blue curves represent respectively the full theory and HEFT predictions where the solid lines are for the NLO QCD distributions while the dotted lines are for the LO ones. In the middle panel the ratio of the full theory over HEFT is displayed. Both at LO and NLO, for $p_{T,h} > 300$ GeV, the crucial region where the BSM distribution separates from the SM, the HEFT approximation overestimates the full theory prediction. Given $\frac{d\sigma^{HEFT}}{dp_{T,h}} / \frac{d\sigma^{full}}{dp_{T,h}}$, the ratio of the full theory distribution over the distribution in the HEFT approximation, at 500 GeV $\frac{d\sigma^{HEFT}}{dp_{T,h}} / \frac{d\sigma^{full}}{dp_{T,h}} \sim 2$, i.e. the magnitude of the HEFT differential cross section is twice the magnitude of the full theory and the ratio increases with $p_{T,h}$. This proves the vital importance of the top quark mass contribution and of the full m_t dependent predictions to study properly new Physics effects in this process.

8.5 Conclusions

We have presented the first NLO QCD calculation with full m_t dependence of Higgs plus one jet production within the EWChL framework. The scattering amplitude of the process depends on the two anomalous couplings c_t and c_{ggh} . We focused on the slice of the anomalous couplings space $(c_t + \frac{3}{2}c_{ggh}) = 1$; couplings belonging to such slice produce total cross sections which are almost the same.

We have presented the total cross sections and differential $p_{T,h}$ distributions for the BSM configuration $(c_t = 0.9, c_{ggh} = \frac{2}{30})$ at LO and NLO QCD and compared them to the SM predictions. We showed that, as at LO, the BSM NLO QCD total cross section with $p_{T,h}^{cut} = 30\text{GeV}$ and the SM one are the extremely similar within the uncertainties. Looking to the Higgs boson transverse momentum distributions, we showed that the two differential cross sections are very close when the Higgs transverse momentum is small and they start to split for $p_{T,h} > 300\text{ GeV}$. This behaviour already known at LO also holds on at NLO QCD. We observed that including BSM effects the differential K-factor is still flat above the top quark threshold, as in the SM case, and its total value is almost the same as the SM highlighting the importance of NLO QCD corrections. We presented the ratio of this BSM total cross section to the SM prediction for different cuts of the transverse momentum, also here finding the similar deviations at LO and NLO QCD. Our analysis has shown that the NLO QCD total and differential cross sections behave as at LO and the high $p_{T,h}$ analysis is crucial to observe New Physics effects.

At last we compared the full top mass dependent and HEFT $p_{T,h}$ distributions at LO and NLO QCD, showing how the HEFT fails to approximate the full theory in the sensible region where SM and BSM contributions disentangle and highlighting the importance of the NLO QCD predictions with full m_t dependency to study correctly this process.

The POWHEG version of the code allowing to vary c_t and c_{ggh} will be publicly available in the POWHEG-BOX-V2 package. This Monte Carlo event generator program could help the experimentalists to match theoretical prediction to observation.

9 Conclusion and Outlook

The Standard Model of particle physics is the theory describing the strong and electroweak interactions of quarks, leptons and bosons. Its predictive power has been intensively tested over the years achieving amazing results. The Higgs sector, the last block of the SM that has been observed, is least explored and precision studies of Higgs parameters are among the main goals of the LHC strategy. Experimental constraints of the Higgs couplings to fermions and other Higgs bosons are still loose, allowing deviations from the SM predictions. Supposing New Particle Physics in the upper TeV range, their effects can be parametrised within an EFT. The EFT can be used to study possible deviations according to the experimental limits.

In this work we studied two processes of interest within such a framework: Higgs boson pair production and Higgs plus one jet production. The former is the most promising process to study the trilinear Higgs coupling but hard to observe at the energy scale reached at colliders, while the latter, observable at current energies, allows to analyze the impact of the anomalous couplings modifications and match them to experimental data. We made our calculation within the framework of the electroweak chiral Lagrangian, a non-linearly realized Effective Field Theory in the Higgs sector.

Regarding the Higgs boson pair production we have calculated the NLO QCD corrections with full m_t dependence where the scattering amplitude is function of five possibly anomalous couplings, c_{hhh} , c_t , c_{tt} , c_{ggh} and c_{gggh} . We gave a parametrisation of the total NLO cross section for Higgs boson pair production in gluon fusion and Higgs boson pair invariant mass distribution in terms of 23 coefficients of all combinations of these couplings and we have shown differential cross sections for m_{hh} and $p_{T,h}$ at several benchmark points which exhibit characteristic shapes of the distributions. We showed the effects of different combinations of the anomalous coupling on the total cross section and on the magnitude and shape of the differential distributions.

Using the parametrization of the total NLO invariant mass distribution, we investigated the possible shapes yielded by certain configurations of the couplings. We visualised the parameter space leading to characteristic shape types following two approaches. In the first approach we defined four characteristic shapes while in the second we used an unsupervised learning algorithm to determine representative shapes. The latter approach resulted to be more efficient to catch the different shape characteristics.

We have calculated the Higgs plus one jet NLO QCD corrections with full m_t dependence. The amplitude of the process is function of the two anomalous couplings c_t and c_{ggh} . In our work we focused on the region $c_t + \frac{3}{2}c_{ggh} = 1$. It has been shown at LO that points belonging to this region yield the same cross section when the cut to $p_{T,h}$ is small. We presented the total cross section and Higgs boson transverse momentum distribution for the SM and the BSM configuration $(c_t, c_{ggh}) = (0.9, \frac{2}{30})$ both lying on

9 Conclusion and Outlook

the slice $c_t + \frac{3}{2}c_{ggh} = 1$. Our analysis displayed how the behaviour, already known at LO, persists when including NLO QCD corrections. We computed the ratios of the BSM configuration to the SM predictions for different cuts of $p_{T,h}$, showing that as at LO as well as at NLO QCD the two cross sections disentangle for high values of $p_{T,h}$, enforcing even more the importance of the high $p_{T,h}$ analysis. The Monte Carlo event generator code will be publicly available.

The last century was characterized by several astonishing discoveries which demonstrated the validity of the SM as the correct framework to describe the fundamental particles and their interactions. With the start of LHC New Physics particles were expected to be found, but until now the only discovery has been the Higgs boson. Now the physics community is wondering when New Physics will finally be observed and if there are still hopes to detect some new particles at LHC. The HL-LHC update will produce a large amount of data and the statistical uncertainty will reduce considerably. This will improve the measurements of the Higgs couplings which are still loosely constrained and maybe, if we are lucky, deviations from the SM predictions could be observed. We hope that the tools we developed and discussed in this work, could help the experimental collaborations in this process of discovery. The observation of New Physics would bring us to a new era of knowledge and understanding of the world. Even if at the moment the Physics community seems to be discouraged about new discoveries the LHC RUN 3, the HL-LHC update and the new particle colliders which will be built in the next years give a reason to hold on and to pursue the research. As Dante wrote "We were not made to live as brutes but to follow virtue and knowledge".

List of Figures

1.1	$pp \rightarrow t\bar{t}H$ total cross sections in proton-proton collisions at centre-mass energy of 8 and 13 TeV. The black dots represent the experimental data including uncertainties. The SM prediction as function of the energy \sqrt{s} is represented by the dashed line, where the band enveloping this line is the theoretical uncertainty [5].	2
3.1	Comparison of the gluon and up quark PDFs at energy scale $Q^2 = 100$ GeV, normalized to the central value of CT14 [76].	14
3.2	$t\bar{t}$ total cross section as function of the invariant mass. The picture shows the experimental data (black vertical lines) and the theoretical predictions (color bars) at different orders [77].	15
3.3	The bubble integral diagram, where p is the momentum entering and exiting the loop and k the loop momentum.	16
4.1	Higgs boson pair production LO diagrams. On the left the diagram of the full theory, on the right the same diagram in the HEFT approximation where the top loop is shrunk.	21
5.1	Perceptron architecture [138].	26
5.2	Architecture of a Fully Connected Network [139, 140].	28
5.3	Autoencoder architecture [138].	29
5.4	Clustering procedure example. The data set is classified in three clusters according to the K-Means criteria [138].	29
6.1	Higgs-pair production in gluon fusion at leading order in the chiral Lagrangian. The black dots indicate vertices from \mathcal{L}_2 , the black squares denote local terms from \mathcal{L}_4	33
6.2	Higgs-pair production in gluon fusion at NLO: Examples for virtual two-loop diagrams at order g_s^4	35
6.3	Higgs-pair production in gluon fusion at NLO: Examples for virtual one-loop diagrams at order g_s^4	36
6.4	Higgs-pair production in gluon fusion at NLO: Tree diagram at order g_s^4	36
6.5	Higgs-pair production in gluon fusion at NLO: Examples for real-emission diagrams at order g_s^3	37
6.6	Higgs-pair production in gluon fusion at NLO: Examples for contributions that are consistently neglected within our approximation. The dotted square indicates a local term at chiral dimension 6 (two-loop order). See text for further explanation.	38

List of Figures

6.7	Iso-contours of σ/σ_{SM} : (a) c_{gghh} and (b) c_{ggh} versus c_{tt}	40
6.8	Iso-contours of σ/σ_{SM} : (a) c_{ggh} versus c_{hhh} and (b) c_{tt} versus c_{hhh}	42
6.9	Iso-contours of σ/σ_{SM} : (a) c_t versus c_{hhh} and (b) c_t versus c_{tt}	43
6.10	Iso-contours of σ/σ_{SM} : (a) c_{gghh} versus c_{ggh} and (b) c_{gghh} versus c_t . . .	44
6.11	K-factors for the total cross section as a function of the different couplings.	45
6.12	Higgs boson pair invariant mass distribution and transverse momentum distribution of one of the Higgs bosons for benchmark point 1, $c_{hhh} = 7.5, c_t = 1, c_{tt} = -1, c_{ggh} = c_{gghh} = 0$. The ratio plot with the K-factor shows NLO_{BSM}/LO_{BSM} . The lower ratio plot shows the ratios (Born-improved NLO HEFT)/ NLO_{BSM} (purple) and FT_{approx}/NLO_{BSM} (green).	45
6.13	Same as Fig. 6.12 but for benchmark point 2, $c_{hhh} = 1, c_t = 1, c_{tt} = 0.5, c_{ggh} = -8/15, c_{gghh} = -0.2$	48
6.14	Same as Fig. 6.12 but for benchmark point 3, $c_{hhh} = 1, c_t = 1, c_{tt} = -1.5, c_{ggh} = 0, c_{gghh} = 4/15$	48
6.15	Same as Fig. 6.12 but for benchmark point 4, $c_{hhh} = -3.5, c_t = 1.5, c_{tt} = -3, c_{ggh} = c_{gghh} = 0$	49
6.16	A point from cluster 4, $c_{hhh} = 1, c_t = 1, c_{tt} = 0, c_{ggh} = 4/15, c_{gghh} = -0.2$, which leads to a similar shape as benchmark point 4, but to a much smaller cross section.	50
6.17	Same as Fig. 6.12 but for benchmark point 5, $c_{hhh} = 1, c_t = 1, c_{tt} = 0, c_{ggh} = 8/15, c_{gghh} = 1/3$	51
6.18	Same as Fig. 6.12 but for benchmark point 6, $c_{hhh} = 2.4, c_t = 1, c_{tt} = 0, c_{ggh} = 2/15, c_{gghh} = 1/15$	51
6.19	Same as Fig. 6.12 but for benchmark point 7, $c_{hhh} = 5, c_t = 1, c_{tt} = 0, c_{ggh} = 2/15, c_{gghh} = 1/15$	52
6.20	Same as Fig. 6.12 but for benchmark point 8a, $c_{hhh} = 1, c_t = 1, c_{tt} = 0.5, c_{ggh} = 4/15, c_{gghh} = 0$	52
6.21	Same as Fig. 6.12 but for benchmark point 9, $c_{hhh} = 1, c_t = 1, c_{tt} = 1, c_{ggh} = -0.4, c_{gghh} = -0.2$	53
6.22	Same as Fig. 6.12 but for benchmark point 10, $c_{hhh} = 10, c_t = 1.5, c_{tt} = -1, c_{ggh} = c_{gghh} = 0$	53
6.23	Same as Fig. 6.12 but for benchmark point 11, $c_{hhh} = 2.4, c_t = 1, c_{tt} = 0, c_{ggh} = 2/3, c_{gghh} = 1/3$	54
6.24	Same as Fig. 6.12 but for benchmark point 12, $c_{hhh} = 15, c_t = 1, c_{tt} = 1, c_{ggh} = c_{gghh} = 0$	54
6.25	Scale variations for benchmark point 5.	54
6.26	Higgs boson pair invariant mass distributions for various values of c_{tt} . .	55
7.1	The four kinds of shapes defined in our analyzer to classify the m_{hh} distributions. The colours correspond to the colours shown in Figs. 7.2 to 7.6.	59

7.2	The parameter regions leading to each predefined shape type in the $c_t - c_{hhh}$ (left) and $c_{ggh} - c_{gghh}$ (right) parameter spaces. The black area denotes shapes of kind 1 (enhanced low m_{hh} region; green: kind 2 (well separated double peaks), red: kind 3 (SM-like), blue: kind 4 (close-by double peaks). The white diamonds mark the Standard Model point. . .	61
7.3	The parameter regions associated to each shape type in the $c_t - c_{tt}$ (left) and $c_{hhh} - c_{tt}$ (right) parameter spaces. For the colour code we refer to Fig. 7.1.	61
7.4	The parameter regions associated to each shape type in the $c_t - c_{ggh}$ (left) and $c_t - c_{gghh}$ (right) planes.	62
7.5	The parameter regions associated to each shape type in the $c_{hhh} - c_{ggh}$ (left) and $c_{hhh} - c_{gghh}$ (right) planes.	62
7.6	The parameter regions associated to each shape type in the $c_{tt} - c_{ggh}$ (left) and $c_{tt} - c_{gghh}$ (right) planes.	63
7.7	Examples of $p_{T,h}$ distributions with a single or double peak.	64
7.8	The parameter space associated to each shape type in the $c_t - c_{hhh}$ and $c_{hhh} - c_{tt}$ planes for the $p_{T,h}$ distribution.	64
7.9	The clusters obtained by asking for a classification into four shape types. We show the cluster centres obtained from 10 different encoder models, in the colour code defined in Table 7.1.	66
7.10	The clusters obtained by asking for a classification into seven shape types. The cluster centres obtained from 10 different encoder models are shown in the colour code defined in Table 7.1.	67
7.11	Shape types produced by variations of c_t versus c_{hhh} (top) and c_t versus c_{tt} (bottom). Left: 4 clusters, right: 7 clusters. The areas outside the silver and white curves are regions where the total cross section exceeds $6.9 \times \sigma_{SM}$ and $22.2 \times \sigma_{SM}$, respectively. These values are motivated by the current ATLAS/CMS limits at $\sqrt{s} = 13$ TeV [32, 33]. The red areas denote SM-like shapes. The full colour code is given in Table 7.1.	69
7.12	Contributions of the coefficients A_i in eq. (7.2) which are relevant for Fig. 7.11.	69
7.13	Shape types produced by variations of c_{hhh} versus c_{tt} (top) and c_{ggh} versus c_{gghh} (bottom). Left: 4 clusters, right: 7 clusters. The areas outside the silver and white curves are regions where the total cross section exceeds $6.9 \times \sigma_{SM}$ and $22.2 \times \sigma_{SM}$, respectively. These values are motivated by the current ATLAS/CMS limits at $\sqrt{s} = 13$ TeV [32, 33]. The red areas denote SM-like shapes. The full colour code is given in Table 7.1.	70
7.14	Shape types produced by variations of c_{hhh} versus c_{ggh} (top) and c_{gghh} (bottom). Left: 4 clusters, right: 7 clusters. The areas outside the silver and white curves are regions where the total cross section exceeds $6.9 \times \sigma_{SM}$ and $22.2 \times \sigma_{SM}$, respectively. These values are motivated by the current ATLAS/CMS limits at $\sqrt{s} = 13$ TeV [32, 33]. The red areas denote SM-like shapes. The full colour code is given in Table 7.1.	71

List of Figures

7.15	Shape types produced by variations of c_{tt} versus c_{ggh} (top) and c_{gghh} (bottom). Left: 4 clusters, right: 7 clusters. The areas outside the silver and white curves are regions where the total cross section exceeds $6.9 \times \sigma_{SM}$ and $22.2 \times \sigma_{SM}$, respectively. These values are motivated by the current ATLAS/CMS limits at $\sqrt{s} = 13$ TeV [32, 33]. The red areas denote SM-like shapes.	72
7.16	Shape types produced by variations of c_t versus c_{ggh} (top) and c_{gghh} (bottom). Left: 4 clusters, right: 7 clusters. The areas outside the silver and white curves are regions where the total cross section exceeds $6.9 \times \sigma_{SM}$ and $22.2 \times \sigma_{SM}$, respectively. These values are motivated by the current ATLAS/CMS limits at $\sqrt{s} = 13$ TeV [32, 33]. The red areas denote SM-like shapes.	73
7.17	Comparison of LO and NLO results for shape types produced by variations of c_t versus c_{hhh} . Left: LO, right: NLO.	74
7.18	Three-dimensional visualisation of shape types produced by variations of c_t , c_{hhh} and c_{ggh} simulating the SMEFT situation. For c_{tt} we used $c_{tt} = 0.05c_t$, while $c_{gghh} = c_{ggh}/2$	75
8.1	Higgs plus one gluon production at leading order in the chiral Lagrangian. The black dots indicate vertices from \mathcal{L}_2 , the black squares denote local terms from \mathcal{L}_4 . We distinguish the light quark and the top quark lines specified by q and t respectively.	79
8.2	Higgs plus one jet production at NLO: Examples for virtual one-loop and two-loop diagrams in the gluon fusion channel.	83
8.3	Higgs plus one jet production at NLO: Examples for real tree-level and one-loop diagrams in the gluon fusion channel.	84
8.4	Higgs boson transverse momentum spectrum at LO and NLO in QCD. The upper panel shows the differential cross sections, in the middle panel we normalize the BSM distributions to the SM predictions and in the lower panel we show the K-factors (the ratio NLO/LO) for both of the distributions.	85
8.5	Higgs boson transverse momentum spectrum at NLO in QCD with full top mass dependence and in the HEFT approximation. The upper panel shows the differential cross sections, in the middle panel we normalize the BSM distributions to the HEFT predictions and in the lower panel we show the K-factors (the ratio NLO/LO) for both the distributions.	87

List of Tables

2.1	Table reassuming the quarks properties. J is the total angular momenta, B the baryon number, Q the electric charge and I_3 the isospin.	6
6.1	Results for the coefficients defined in Eq. (6.4). The uncertainties are obtained from the uncertainties on the total cross sections entering the projections, using error propagation which neglects correlations between these cross sections.	41
6.2	Benchmark points used for the distributions shown below.	46
6.3	Translation between the conventions for the definition of the anomalous couplings.	46
6.4	Total cross sections at NLO (second column) including the K-factor (third column), scale uncertainties (4th column) and statistical uncertainties (5th column) and the ratio to the SM total NLO cross section (6th column).	47
7.1	Clusters and shape types with corresponding colour codes for the classification into four and seven clusters.	68
8.1	LO and NLO total cross section with statistical error for $c_t = 0.9$, $c_{ggh} = \frac{2}{30}$ and the SM for $p_{T,h} > 30$ GeV. The upper and lower values due to scale variation are shown.	84
8.2	LO and NLO total cross section for $(c_t = 0.9, c_{ggh} = \frac{2}{30})$ normalised to the SM value for different $p_{T,h}^{cut}$, including the statistical error.	86

Bibliography

- [1] ATLAS collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett. B* **716** (2012) 1 [1207.7214].
- [2] CMS collaboration, *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*, *Phys. Lett. B* **716** (2012) 30 [1207.7235].
- [3] PARTICLE DATA GROUP collaboration, *Review of Particle Physics*, *Phys. Rev. D* **98** (2018) 030001.
- [4] J. Aebischer, W. Altmannshofer, D. Guadagnoli, M. Reboud, P. Stangl and D. M. Straub, *B-decay discrepancies after Moriond 2019*, 1903.10434.
- [5] ATLAS collaboration, *Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector*, *Phys. Lett. B* **784** (2018) 173 [1806.00425].
- [6] C.-N. Yang and R. L. Mills, *Conservation of Isotopic Spin and Isotopic Gauge Invariance*, *Phys. Rev.* **96** (1954) 191.
- [7] M. Gell-Mann, *A Schematic Model of Baryons and Mesons*, *Phys. Lett.* **8** (1964) 214.
- [8] G. Zweig, *An $SU(3)$ model for strong interaction symmetry and its breaking. Version 2*, in *DEVELOPMENTS IN THE QUARK THEORY OF HADRONS. VOL. 1. 1964 - 1978*, D. Lichtenberg and S. P. Rosen, eds., pp. 22–101, (1964).
- [9] S. L. Glashow, J. Iliopoulos and L. Maiani, *Weak interactions with lepton-hadron symmetry*, *Phys. Rev. D* **2** (1970) 1285.
- [10] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, *Phys. Rev. Lett.* **10** (1963) 531.
- [11] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, *Prog. Theor. Phys.* **49** (1973) 652.
- [12] E. D. Bloom et al., *High-Energy Inelastic $e p$ Scattering at 6-Degrees and 10-Degrees*, *Phys. Rev. Lett.* **23** (1969) 930.
- [13] M. Breidenbach, J. I. Friedman, H. W. Kendall, E. D. Bloom, D. H. Coward, H. C. DeStaebler et al., *Observed Behavior of Highly Inelastic electron-Proton Scattering*, *Phys. Rev. Lett.* **23** (1969) 935.

Bibliography

- [14] SLAC-SP-017 collaboration, *Discovery of a Narrow Resonance in e^+e^- Annihilation*, *Phys. Rev. Lett.* **33** (1974) 1406.
- [15] E598 collaboration, *Experimental Observation of a Heavy Particle J* , *Phys. Rev. Lett.* **33** (1974) 1404.
- [16] S. W. Herb et al., *Observation of a Dimuon Resonance at 9.5-GeV in 400-GeV Proton-Nucleus Collisions*, *Phys. Rev. Lett.* **39** (1977) 252.
- [17] B. R. Stella and H.-J. Meyer, *$Y(9.46 \text{ GeV})$ and the gluon discovery (a critical recollection of PLUTO results)*, *Eur. Phys. J.* **H36** (2011) 203 [1008.1869].
- [18] CDF collaboration, *Observation of top quark production in $\bar{p}p$ collisions*, *Phys. Rev. Lett.* **74** (1995) 2626 [hep-ex/9503002].
- [19] D0 collaboration, *Search for high mass top quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.8 \text{ TeV}$* , *Phys. Rev. Lett.* **74** (1995) 2422 [hep-ex/9411001].
- [20] S. Weinberg, *A Model of Leptons*, *Phys. Rev. Lett.* **19** (1967) 1264.
- [21] A. Salam, *Weak and Electromagnetic Interactions*, *Conf. Proc.* **C680519** (1968) 367.
- [22] S. L. Glashow, *The renormalizability of vector meson interactions*, *Nucl. Phys.* **10** (1959) 107.
- [23] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, *Phys. Rev. Lett.* **13** (1964) 321.
- [24] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, *Phys. Rev. Lett.* **13** (1964) 508.
- [25] G. S. Guralnik, C. R. Hagen and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, *Phys. Rev. Lett.* **13** (1964) 585.
- [26] UA1 collaboration, *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $s^{1/2} = 540\text{-GeV}$* , *Phys. Lett.* **122B** (1983) 103.
- [27] UA2 collaboration, *Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti- p p Collider*, *Phys. Lett.* **122B** (1983) 476.
- [28] UA1 collaboration, *Experimental Observation of Lepton Pairs of Invariant Mass Around $95\text{-GeV}/c^2$ at the CERN SPS Collider*, *Phys. Lett.* **126B** (1983) 398.
- [29] UA2 collaboration, *Evidence for $Z^0 \rightarrow e^+e^-$ at the CERN anti- p p Collider*, *Phys. Lett.* **129B** (1983) 130.

- [30] CMS collaboration, *Observation of $t\bar{t}H$ production*, *Phys. Rev. Lett.* **120** (2018) 231801 [1804.02610].
- [31] M. Capozzi and G. Heinrich, *Exploring anomalous couplings in Higgs boson pair production through shape analysis*, 1908.08923.
- [32] ATLAS collaboration, *Combination of searches for Higgs boson pairs in pp collisions at $\sqrt{s}=13$ TeV with the ATLAS detector*, 1906.02025.
- [33] CMS collaboration, *Combination of searches for Higgs boson pair production in proton-proton collisions at $\sqrt{s}=13$ TeV*, *Phys. Rev. Lett.* **122** (2019) 121803 [1811.09689].
- [34] M. McCullough, *An Indirect Model-Dependent Probe of the Higgs Self-Coupling*, *Phys. Rev.* **D90** (2014) 015001 [1312.3322].
- [35] M. Gorbahn and U. Haisch, *Indirect probes of the trilinear Higgs coupling: $gg \rightarrow h$ and $h \rightarrow \gamma\gamma$* , *JHEP* **10** (2016) 094 [1607.03773].
- [36] G. Degrandi, P. P. Giardino, F. Maltoni and D. Pagani, *Probing the Higgs self coupling via single Higgs production at the LHC*, *JHEP* **12** (2016) 080 [1607.04251].
- [37] W. Bizon, M. Gorbahn, U. Haisch and G. Zanderighi, *Constraints on the trilinear Higgs coupling from vector boson fusion and associated Higgs production at the LHC*, *JHEP* **07** (2017) 083 [1610.05771].
- [38] F. Maltoni, D. Pagani, A. Shivaji and X. Zhao, *Trilinear Higgs coupling determination via single-Higgs differential measurements at the LHC*, *Eur. Phys. J.* **C77** (2017) 887 [1709.08649].
- [39] G. D. Kribs, A. Maier, H. Rzehak, M. Spannowsky and P. Waite, *Electroweak oblique parameters as a probe of the trilinear Higgs boson self-interaction*, *Phys. Rev.* **D95** (2017) 093004 [1702.07678].
- [40] G. Degrandi, M. Fedele and P. P. Giardino, *Constraints on the trilinear Higgs self coupling from precision observables*, *JHEP* **04** (2017) 155 [1702.01737].
- [41] J. Nakamura and A. Shivaji, *Direct measurement of the Higgs self-coupling in $e^+e^- \rightarrow ZH$* , 1812.01576.
- [42] W. Kilian, S. Sun, Q.-S. Yan, X. Zhao and Z. Zhao, *Multi-Higgs Production and Unitarity in Vector-Boson Fusion at Future Hadron Colliders*, 1808.05534.
- [43] F. Maltoni, D. Pagani and X. Zhao, *Constraining the Higgs self-couplings at e^+e^- colliders*, *JHEP* **07** (2018) 087 [1802.07616].
- [44] E. Vryonidou and C. Zhang, *Dimension-six electroweak top-loop effects in Higgs production and decay*, *JHEP* **08** (2018) 036 [1804.09766].

Bibliography

- [45] M. Gorbahn and U. Haisch, *Two-loop amplitudes for Higgs plus jet production involving a modified trilinear Higgs coupling*, *JHEP* **04** (2019) 062 [1902.05480].
- [46] ATLAS COLLABORATION collaboration, *Constraint of the Higgs boson self-coupling from Higgs boson differential production and decay measurements*, Tech. Rep. ATL-PHYS-PUB-2019-009, CERN, Geneva, Mar, 2019.
- [47] ATLAS collaboration, *Constraints on the Higgs boson self-coupling from the combination of single-Higgs and double-Higgs production analyses performed with the ATLAS experiment*, tech. rep., 2019.
- [48] W. Bizon, U. Haisch and L. Rottoli, *Constraints on the quartic Higgs self-coupling from double-Higgs production at future hadron colliders*, 1810.04665.
- [49] S. Borowka, C. Duhr, F. Maltoni, D. Pagani, A. Shivaji and X. Zhao, *Probing the scalar potential via double Higgs boson production at hadron colliders*, *JHEP* **04** (2019) 016 [1811.12366].
- [50] A. Falkowski and R. Rattazzi, *Which EFT*, 1902.05936.
- [51] S. Chang and M. A. Luty, *The Higgs Trilinear Coupling and the Scale of New Physics*, 1902.05556.
- [52] L. Di Luzio, R. Gröber and M. Spannowsky, *Maxi-sizing the trilinear Higgs self-coupling: how large could it be?*, *Eur. Phys. J.* **C77** (2017) 788 [1704.02311].
- [53] S. Di Vita, C. Grojean, G. Panico, M. Riembau and T. Vantalon, *A global view on the Higgs self-coupling*, *JHEP* **09** (2017) 069 [1704.01953].
- [54] J. Braathen and S. Kanemura, *On two-loop corrections to the Higgs trilinear coupling in models with extended scalar sectors*, *Phys. Lett.* **B796** (2019) 38 [1903.05417].
- [55] P. Basler, S. Dawson, C. Englert and M. Mühlleitner, *Showcasing HH production: Benchmarks for the LHC and HL-LHC*, *Phys. Rev.* **D99** (2019) 055048 [1812.03542].
- [56] K. S. Babu and S. Jana, *Enhanced Di-Higgs Production in the Two Higgs Doublet Model*, *JHEP* **02** (2019) 193 [1812.11943].
- [57] A. Adhikary, S. Banerjee, R. K. Barman, B. Bhattacharjee and S. Niyogi, *Revisiting the non-resonant Higgs pair production at the HL-LHC*, *Physics* **2018** (2018) 116 [1712.05346].
- [58] I. M. Lewis and M. Sullivan, *Benchmarks for Double Higgs Production in the Singlet Extended Standard Model at the LHC*, *Phys. Rev.* **D96** (2017) 035037 [1701.08774].

- [59] S. Dawson, C. Englert and T. Plehn, *Higgs Physics: It ain't over till it's over*, *Phys. Rept.* **816** (2019) 1 [1808.01324].
- [60] M. Cepeda et al., *Higgs Physics at the HL-LHC and HE-LHC, HL/HE WG2 group report*, 1902.00134.
- [61] J. Alison et al., *Higgs Boson Pair Production at Colliders: Status and Perspectives*, in *Double Higgs Production at Colliders Batavia, IL, USA, September 4, 2018-9, 2019*, B. Di Micco, M. Gouzevitch, J. Mazzitelli and C. Vernieri, eds., 2019, 1910.00012, <https://lss.fnal.gov/archive/2019/conf/fermilab-conf-19-468-e-t.pdf>.
- [62] J. Baglio, A. Djouadi and J. Quevillon, *Prospects for Higgs physics at energies up to 100 TeV*, *Rept. Prog. Phys.* **79** (2016) 116201 [1511.07853].
- [63] K. G. Chetyrkin and M. F. Zoller, *Three-loop beta-functions for top-Yukawa and the Higgs self-interaction in the Standard Model*, *JHEP* **06** (2012) 033 [1205.2892].
- [64] W. Hollik, *Quantum field theory and the Standard Model*, in *High-energy physics. Proceedings, 17th European School, ESHEP 2009, Bautzen, Germany, June 14-27, 2009*, 2010, 1012.3883.
- [65] G. Heinrich, “*Colourful Loops: Introduction to Quantum Chromodynamics and Loop Calculations*,” Technische Universität München, München.” <http://wwwth.mpp.mpg.de/members/gudrun>, 2018.
- [66] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*. Addison-Wesley, Reading, USA, 1995.
- [67] S. Weinberg, *The Quantum theory of fields. Vol. 1: Foundations*. Cambridge University Press, 2005.
- [68] O. Benhar, “*Introduction to QED*,” University of Rome La Sapienza, Rome.” <https://chimera.roma1.infn.it/OMAR/QED>, 2017.
- [69] J. C. Collins, D. E. Soper and G. F. Sterman, *Factorization of Hard Processes in QCD*, *Adv. Ser. Direct. High Energy Phys.* **5** (1989) 1 [hep-ph/0409313].
- [70] V. N. Gribov and L. N. Lipatov, *Deep inelastic e p scattering in perturbation theory*, *Sov. J. Nucl. Phys.* **15** (1972) 438.
- [71] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.*, *Sov. Phys. JETP* **46** (1977) 641.
- [72] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, *Nucl. Phys.* **B126** (1977) 298.

Bibliography

- [73] S. Dulat, T.-J. Hou, J. Gao, M. Guzzi, J. Huston, P. Nadolsky et al., *New parton distribution functions from a global analysis of quantum chromodynamics*, *Phys. Rev.* **D93** (2016) 033006 [1506.07443].
- [74] L. A. Harland-Lang, A. D. Martin, P. Motylinski and R. S. Thorne, *Parton distributions in the LHC era: MMHT 2014 PDFs*, *Eur. Phys. J.* **C75** (2015) 204 [1412.3989].
- [75] NNPDF collaboration, *Parton distributions for the LHC Run II*, *JHEP* **04** (2015) 040 [1410.8849].
- [76] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, *J. Phys.* **G43** (2016) 023001 [1510.03865].
- [77] P. Fiedler, *Top pair cross section at NNLO*, in *Proceedings, 49th Rencontres de Moriond on QCD and High Energy Interactions: La Thuile, Italy, March 22-29, 2014*, pp. 161–164, 2014.
- [78] G. 't Hooft and M. J. G. Veltman, *Regularization and Renormalization of Gauge Fields*, *Nucl. Phys.* **B44** (1972) 189.
- [79] G. 't Hooft and M. J. G. Veltman, *Scalar One Loop Integrals*, *Nucl. Phys.* **B153** (1979) 365.
- [80] H. D. Politzer, *Reliable Perturbative Results for Strong Interactions?*, *Phys. Rev. Lett.* **30** (1973) 1346.
- [81] D. J. Gross and F. Wilczek, *Asymptotically Free Gauge Theories - I*, *Phys. Rev.* **D8** (1973) 3633.
- [82] T. Kinoshita, *Mass singularities of Feynman amplitudes*, *J. Math. Phys.* **3** (1962) 650.
- [83] T. D. Lee and M. Nauenberg, *Degenerate Systems and Mass Singularities*, *Phys. Rev.* **133** (1964) B1549.
- [84] S. Catani and M. H. Seymour, *A General algorithm for calculating jet cross-sections in NLO QCD*, *Nucl. Phys.* **B485** (1997) 291 [hep-ph/9605323].
- [85] W. Buchmüller and D. Wyler, *Effective Lagrangian Analysis of New Interactions and Flavor Conservation*, *Nucl. Phys.* **B268** (1986) 621.
- [86] B. Grzadkowski, M. Iskrzynski, M. Misiak and J. Rosiek, *Dimension-Six Terms in the Standard Model Lagrangian*, *JHEP* **10** (2010) 085 [1008.4884].
- [87] L. Berthier and M. Trott, *Towards consistent Electroweak Precision Data constraints in the SMEFT*, *JHEP* **05** (2015) 024 [1502.02570].
- [88] M. Ghezzi, R. Gomez-Ambrosio, G. Passarino and S. Uccirati, *NLO Higgs effective field theory and kappa-framework*, *JHEP* **07** (2015) 175 [1505.03706].

- [89] J. de Blas, M. Ciuchini, E. Franco, S. Mishima, M. Pierini, L. Reina et al., *The Global Electroweak and Higgs Fits in the LHC era*, *PoS EPS-HEP2017* (2017) 467 [1710.05402].
- [90] I. Brivio and M. Trott, *Scheming in the SMEFT... and a reparameterization invariance!*, *JHEP* **07** (2017) 148 [1701.06424].
- [91] J. Ellis, C. W. Murphy, V. Sanz and T. You, *Updated Global SMEFT Fit to Higgs, Diboson and Electroweak Data*, 1803.03252.
- [92] F. Feruglio, *The Chiral approach to the electroweak interactions*, *Int. J. Mod. Phys. A* **8** (1993) 4937 [hep-ph/9301281].
- [93] J. Bagger, V. D. Barger, K.-m. Cheung, J. F. Gunion, T. Han, G. A. Ladinsky et al., *The Strongly interacting $W W$ system: Gold plated modes*, *Phys. Rev. D* **49** (1994) 1246 [hep-ph/9306256].
- [94] V. Koulovassilopoulos and R. S. Chivukula, *The Phenomenology of a nonstandard Higgs boson in $W(L) W(L)$ scattering*, *Phys. Rev. D* **50** (1994) 3218 [hep-ph/9312317].
- [95] C. P. Burgess, J. Matias and M. Pospelov, *A Higgs or not a Higgs? What to do if you discover a new scalar particle*, *Int. J. Mod. Phys. A* **17** (2002) 1841 [hep-ph/9912459].
- [96] L.-M. Wang and Q. Wang, *Electroweak chiral Lagrangian for neutral Higgs boson*, *Chin. Phys. Lett.* **25** (2008) 1984 [hep-ph/0605104].
- [97] B. Grinstein and M. Trott, *A Higgs-Higgs bound state due to new physics at a TeV*, *Phys. Rev. D* **76** (2007) 073002 [0704.1505].
- [98] R. Contino, C. Grojean, M. Moretti, F. Piccinini and R. Rattazzi, *Strong Double Higgs Production at the LHC*, *JHEP* **05** (2010) 089 [1002.1011].
- [99] R. Contino, *The Higgs as a Composite Nambu-Goldstone Boson*, in *Physics of the large and the small, TASI 09, proceedings of the Theoretical Advanced Study Institute in Elementary Particle Physics, Boulder, Colorado, USA, 1-26 June 2009*, pp. 235–306, 2011, 1005.4269, DOI.
- [100] R. Alonso, M. B. Gavela, L. Merlo, S. Rigolin and J. Yepes, *The Effective Chiral Lagrangian for a Light Dynamical "Higgs Particle"*, *Phys. Lett. B* **722** (2013) 330 [1212.3305].
- [101] G. Buchalla, O. Catà and C. Krause, *Complete Electroweak Chiral Lagrangian with a Light Higgs at NLO*, *Nucl. Phys. B* **880** (2014) 552 [1307.5017].
- [102] R. L. Delgado, A. Dobado and F. J. Llanes-Estrada, *One-loop $W_L W_L$ and $Z_L Z_L$ scattering from the electroweak Chiral Lagrangian with a light Higgs-like scalar*, *JHEP* **02** (2014) 121 [1311.5993].

Bibliography

- [103] G. Buchalla, O. Catá and C. Krause, *On the Power Counting in Effective Field Theories*, *Phys. Lett.* **B731** (2014) 80 [1312.5624].
- [104] G. Buchalla, O. Cata, A. Celis and C. Krause, *Fitting Higgs Data with Nonlinear Effective Theory*, *Eur. Phys. J.* **C76** (2016) 233 [1511.00988].
- [105] J. de Blas, O. Eberhardt and C. Krause, *Current and future constraints on Higgs couplings in the nonlinear Effective Theory*, 1803.00939.
- [106] LHC HIGGS CROSS SECTION WORKING GROUP collaboration, *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, 1610.07922.
- [107] S. Weinberg, *Baryon and Lepton Nonconserving Processes*, *Phys. Rev. Lett.* **43** (1979) 1566.
- [108] A. de Gouvea, J. Herrero-Garcia and A. Kobach, *Neutrino Masses, Grand Unification, and Baryon Number Violation*, *Phys. Rev.* **D90** (2014) 016011 [1404.4057].
- [109] F. F. Deppisch, J. Harz, M. Hirsch, W.-C. Huang and H. Päs, *Falsifying High-Scale Baryogenesis with Neutrinoless Double Beta Decay and Lepton Flavor Violation*, *Phys. Rev.* **D92** (2015) 036005 [1503.04825].
- [110] G. Buchalla, O. Cata, A. Celis, M. Knecht and C. Krause, *Complete One-Loop Renormalization of the Higgs-Electroweak Chiral Lagrangian*, *Nucl. Phys.* **B928** (2018) 93 [1710.06412].
- [111] G. Buchalla, M. Capozzi, A. Celis, G. Heinrich and L. Scyboz, *Higgs boson pair production in non-linear Effective Field Theory with full m_t -dependence at NLO QCD*, *JHEP* **09** (2018) 057 [1806.05162].
- [112] G. Buchalla, O. Cata, A. Celis and C. Krause, *Note on Anomalous Higgs-Boson Couplings in Effective Field Theory*, *Phys. Lett.* **B750** (2015) 298 [1504.01707].
- [113] L. de Oliveira, M. Kagan, L. Mackey, B. Nachman and A. Schwartzman, *Jet-images – deep learning edition*, *JHEP* **07** (2016) 069 [1511.05190].
- [114] J. Brehmer, K. Cranmer, F. Kling and T. Plehn, *Better Higgs boson measurements through information geometry*, *Phys. Rev.* **D95** (2017) 073002 [1612.05261].
- [115] J. Brehmer, K. Cranmer, G. Louppe and J. Pavez, *Constraining Effective Field Theories with Machine Learning*, *Phys. Rev. Lett.* **121** (2018) 111801 [1805.00013].
- [116] J. Brehmer, K. Cranmer, G. Louppe and J. Pavez, *A Guide to Constraining Effective Field Theories with Machine Learning*, *Phys. Rev.* **D98** (2018) 052004 [1805.00020].

- [117] D. Guest, J. Collado, P. Baldi, S.-C. Hsu, G. Urban and D. Whiteson, *Jet Flavor Classification in High-Energy Physics with Deep Neural Networks*, *Phys. Rev. D* **94** (2016) 112002 [1607.08633].
- [118] G. Kasieczka, T. Plehn, M. Russell and T. Schell, *Deep-learning Top Taggers or The End of QCD?*, *JHEP* **05** (2017) 006 [1701.08784].
- [119] K. Datta and A. J. Larkoski, *Novel Jet Observables from Machine Learning*, *JHEP* **03** (2018) 086 [1710.01305].
- [120] G. Louppe, K. Cho, C. Becot and K. Cranmer, *QCD-Aware Recursive Neural Networks for Jet Physics*, *JHEP* **01** (2019) 057 [1702.00748].
- [121] A. J. Larkoski, I. Moult and B. Nachman, *Jet Substructure at the Large Hadron Collider: A Review of Recent Advances in Theory and Machine Learning*, 1709.04464.
- [122] S. Macaluso and D. Shih, *Pulling Out All the Tops with Computer Vision and Deep Learning*, *JHEP* **10** (2018) 121 [1803.00107].
- [123] S. Bollweg, M. Haussmann, G. Kasieczka, M. Luchmann, T. Plehn and J. Thompson, *Deep-Learning Jets with Uncertainties and More*, 1904.10004.
- [124] A. Butter et al., *The Machine Learning Landscape of Top Taggers*, 1902.09914.
- [125] E. A. Moreno, O. Cerri, J. M. Duarte, H. B. Newman, T. Q. Nguyen, A. Periwai et al., *JEDI-net: a jet identification algorithm based on interaction networks*, 1908.05318.
- [126] Y.-C. J. Chen, C.-W. Chiang, G. Cottin and D. Shih, *Boosted W/Z Tagging with Jet Charge and Deep Learning*, 1908.08256.
- [127] S. Chang, T. Cohen and B. Ostdiek, *What is the Machine Learning?*, *Phys. Rev. D* **97** (2018) 056009 [1709.10106].
- [128] J. Hajer, Y.-Y. Li, T. Liu and H. Wang, *Novelty Detection Meets Collider Physics*, 1807.10261.
- [129] C. Englert, P. Galler, P. Harris and M. Spannowsky, *Machine Learning Uncertainties with Adversarial Neural Networks*, *Eur. Phys. J. C* **79** (2019) 4 [1807.08763].
- [130] A. De Simone and T. Jacques, *Guiding New Physics Searches with Unsupervised Learning*, *Eur. Phys. J. C* **79** (2019) 289 [1807.06038].
- [131] A. Andreassen, I. Feige, C. Frye and M. D. Schwartz, *JUNIPR: a Framework for Unsupervised Machine Learning in Particle Physics*, *Eur. Phys. J. C* **79** (2019) 102 [1804.09720].

Bibliography

- [132] R. T. D'Agnolo and A. Wulzer, *Learning New Physics from a Machine*, *Phys. Rev.* **D99** (2019) 015014 [1806.02350].
- [133] A. Blance, M. Spannowsky and P. Waite, *Adversarially-trained autoencoders for robust unsupervised new physics searches*, 1905.10384.
- [134] A. Butter, T. Plehn and R. Winterhalder, *How to GAN LHC Events*, 1907.03764.
- [135] J. Chang, K. Cheung, J. S. Lee and J. Park, *Probing Trilinear Higgs Self-coupling at the HL-LHC via Multivariate Analysis*, 1908.00753.
- [136] J. Brehmer, S. Dawson, S. Homiller, F. Kling and T. Plehn, *Benchmarking simplified template cross sections in WH production*, 1908.06980.
- [137] S. Carrazza and J. Cruz-Martinez, *Towards a new generation of parton densities with deep learning models*, 1907.05075.
- [138] <https://towardsdatascience.com>.
- [139] <http://www.astroml.org>.
- [140] <http://www.datastuff.tech>.
- [141] ATLAS, CMS collaboration, *Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV*, *JHEP* **08** (2016) 045 [1606.02266].
- [142] A. Azatov, R. Contino, G. Panico and M. Son, *Effective field theory analysis of double Higgs boson production via gluon fusion*, *Phys. Rev.* **D92** (2015) 035001 [1502.00539].
- [143] A. Carvalho, M. Dall'Osso, T. Dorigo, F. Goertz, C. A. Gottardo and M. Tosi, *Higgs Pair Production: Choosing Benchmarks With Cluster Analysis*, *JHEP* **04** (2016) 126 [1507.02245].
- [144] A. Carvalho, M. Dall'Osso, P. De Castro Manzano, T. Dorigo, F. Goertz, M. Gouzevich et al., *Analytical parametrization and shape classification of anomalous HH production in the EFT approach*, 1608.06578.
- [145] A. Carvalho, F. Goertz, K. Mimasu, M. Gouzevitch and A. Aggarwal, *On the reinterpretation of non-resonant searches for Higgs boson pairs*, 1710.08261.
- [146] A. Pierce, J. Thaler and L.-T. Wang, *Disentangling Dimension Six Operators through Di-Higgs Boson Production*, *JHEP* **05** (2007) 070 [hep-ph/0609049].
- [147] R. Contino, M. Ghezzi, M. Moretti, G. Panico, F. Piccinini and A. Wulzer, *Anomalous Couplings in Double Higgs Production*, *JHEP* **08** (2012) 154 [1205.5444].

- [148] J. Baglio, A. Djouadi, R. Gröber, M. M. Mühlleitner, J. Quevillon and M. Spira, *The measurement of the Higgs self-coupling at the LHC: theoretical status*, *JHEP* **04** (2013) 151 [1212.5581].
- [149] S. Dawson, E. Furlan and I. Lewis, *Unravelling an extended quark sector through multiple Higgs production?*, *Phys. Rev.* **D87** (2013) 014007 [1210.6663].
- [150] M. J. Dolan, C. Englert and M. Spannowsky, *Higgs self-coupling measurements at the LHC*, *JHEP* **10** (2012) 112 [1206.5001].
- [151] F. Goertz, A. Papaefstathiou, L. L. Yang and J. Zurita, *Higgs boson pair production in the $D=6$ extension of the SM*, *JHEP* **04** (2015) 167 [1410.3471].
- [152] A. J. Barr, M. J. Dolan, C. Englert, D. E. Ferreira de Lima and M. Spannowsky, *Higgs Self-Coupling Measurements at a 100 TeV Hadron Collider*, *JHEP* **02** (2015) 016 [1412.7154].
- [153] M. J. Dolan, C. Englert, N. Greiner, K. Nordstrom and M. Spannowsky, *hhjj production at the LHC*, *Eur. Phys. J.* **C75** (2015) 387 [1506.08008].
- [154] J. K. Behr, D. Bortoletto, J. A. Frost, N. P. Hartland, C. Issever and J. Rojo, *Boosting Higgs pair production in the $b\bar{b}b\bar{b}$ final state with multivariate techniques*, *Eur. Phys. J.* **C76** (2016) 386 [1512.08928].
- [155] F. Maltoni, E. Vryonidou and C. Zhang, *Higgs production in association with a top-antitop pair in the Standard Model Effective Field Theory at NLO in QCD*, *JHEP* **10** (2016) 123 [1607.05330].
- [156] F. Kling, T. Plehn and P. Schichtel, *Maximizing the significance in Higgs boson pair analyses*, *Phys. Rev.* **D95** (2017) 035026 [1607.07441].
- [157] Q.-H. Cao, G. Li, B. Yan, D.-M. Zhang and H. Zhang, *Double Higgs production at the 14 TeV LHC and a 100 TeV pp collider*, *Phys. Rev.* **D96** (2017) 095031 [1611.09336].
- [158] T. Corbett, A. Joglekar, H.-L. Li and J.-H. Yu, *Exploring Extended Scalar Sectors with Di-Higgs Signals: A Higgs EFT Perspective*, *JHEP* **05** (2018) 061 [1705.02551].
- [159] S. Dawson and C. W. Murphy, *Standard Model EFT and Extended Scalar Sectors*, *Phys. Rev.* **D96** (2017) 015041 [1704.07851].
- [160] A. Alves, T. Ghosh and K. Sinha, *Can We Discover Double Higgs Production at the LHC?*, *Phys. Rev.* **D96** (2017) 035022 [1704.07395].
- [161] J. H. Kim, Y. Sakaki and M. Son, *Combined analysis of double Higgs production via gluon fusion at the HL-LHC in the effective field theory approach*, 1801.06093.

Bibliography

- [162] D. Gonçalves, T. Han, F. Kling, T. Plehn and M. Takeuchi, *Higgs Pair Production at Future Hadron Colliders: From Kinematics to Dynamics*, 1802.04319.
- [163] O. J. P. Eboli, G. C. Marques, S. F. Novaes and A. A. Natale, *Twin Higgs boson production*, *Phys. Lett.* **B197** (1987) 269.
- [164] E. W. N. Glover and J. J. van der Bij, *HIGGS BOSON PAIR PRODUCTION VIA GLUON FUSION*, *Nucl. Phys.* **B309** (1988) 282.
- [165] T. Plehn, M. Spira and P. M. Zerwas, *Pair production of neutral Higgs particles in gluon-gluon collisions*, *Nucl. Phys.* **B479** (1996) 46 [hep-ph/9603205].
- [166] S. Dawson, S. Dittmaier and M. Spira, *Neutral Higgs boson pair production at hadron colliders: QCD corrections*, *Phys. Rev.* **D58** (1998) 115012 [hep-ph/9805244].
- [167] R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, P. Torrielli et al., *Higgs pair production at the LHC with NLO and parton-shower effects*, *Phys. Lett.* **B732** (2014) 142 [1401.7340].
- [168] F. Maltoni, E. Vryonidou and M. Zaro, *Top-quark mass effects in double and triple Higgs production in gluon-gluon fusion at NLO*, *JHEP* **11** (2014) 079 [1408.6542].
- [169] J. Grigo, J. Hoff, K. Melnikov and M. Steinhauser, *On the Higgs boson pair production at the LHC*, *Nucl. Phys.* **B875** (2013) 1 [1305.7340].
- [170] J. Grigo, K. Melnikov and M. Steinhauser, *Virtual corrections to Higgs boson pair production in the large top quark mass limit*, *Nucl. Phys.* **B888** (2014) 17 [1408.2422].
- [171] J. Grigo, J. Hoff and M. Steinhauser, *Higgs boson pair production: top quark mass effects at NLO and NNLO*, *Nucl. Phys.* **B900** (2015) 412 [1508.00909].
- [172] G. Degrandi, P. P. Giardino and R. Gröber, *On the two-loop virtual QCD corrections to Higgs boson pair production in the Standard Model*, *Eur. Phys. J.* **C76** (2016) 411 [1603.00385].
- [173] D. de Florian and J. Mazzitelli, *Two-loop virtual corrections to Higgs pair production*, *Phys. Lett.* **B724** (2013) 306 [1305.5206].
- [174] D. de Florian and J. Mazzitelli, *Higgs Boson Pair Production at Next-to-Next-to-Leading Order in QCD*, *Phys. Rev. Lett.* **111** (2013) 201801 [1309.6594].
- [175] D. de Florian, M. Grazzini, C. Hanga, S. Kallweit, J. M. Lindert, P. Maierhöfer et al., *Differential Higgs Boson Pair Production at Next-to-Next-to-Leading Order in QCD*, *JHEP* **09** (2016) 151 [1606.09519].

- [176] D. Y. Shao, C. S. Li, H. T. Li and J. Wang, *Threshold resummation effects in Higgs boson pair production at the LHC*, *JHEP* **07** (2013) 169 [1301.1245].
- [177] D. de Florian and J. Mazzitelli, *Higgs pair production at next-to-next-to-leading logarithmic accuracy at the LHC*, *JHEP* **09** (2015) 053 [1505.07122].
- [178] L.-B. Chen, H. T. Li, H.-S. Shao and J. Wang, *The gluon-fusion production of Higgs boson pair: N^3LO QCD corrections and top-quark mass effects*, 1912.13001.
- [179] S. Borowka, N. Greiner, G. Heinrich, S. P. Jones, M. Kerner, J. Schlenk et al., *Higgs Boson Pair Production in Gluon Fusion at Next-to-Leading Order with Full Top-Quark Mass Dependence*, *Phys. Rev. Lett.* **117** (2016) 012001 [1604.06447].
- [180] S. Borowka, N. Greiner, G. Heinrich, S. P. Jones, M. Kerner, J. Schlenk et al., *Full top quark mass dependence in Higgs boson pair production at NLO*, *JHEP* **10** (2016) 107 [1608.04798].
- [181] G. Ferrera and J. Pires, *Transverse-momentum resummation for Higgs boson pair production at the LHC with top-quark mass effects*, *JHEP* **02** (2017) 139 [1609.01691].
- [182] G. Heinrich, S. P. Jones, M. Kerner, G. Luisoni and E. Vryonidou, *NLO predictions for Higgs boson pair production with full top quark mass dependence matched to parton showers*, *JHEP* **08** (2017) 088 [1703.09252].
- [183] S. Jones and S. Kuttimalai, *Parton Shower and NLO-Matching uncertainties in Higgs Boson Pair Production*, *JHEP* **02** (2018) 176 [1711.03319].
- [184] G. Heinrich, S. P. Jones, M. Kerner, G. Luisoni and L. Scyboz, *Probing the trilinear Higgs boson coupling in di-Higgs production at NLO QCD including parton shower effects*, *JHEP* **06** (2019) 066 [1903.08137].
- [185] J. Davies, G. Heinrich, S. P. Jones, M. Kerner, G. Mishima, M. Steinhauser et al., *Double Higgs boson production at NLO: combining the exact numerical result and high-energy expansion*, *JHEP* **11** (2019) 024 [1907.06408].
- [186] J. Baglio, F. Campanario, S. Glaus, M. Mühlleitner, M. Spira and J. Streicher, *Gluon fusion into Higgs pairs at NLO QCD and the top mass scheme*, *Eur. Phys. J.* **C79** (2019) 459 [1811.05692].
- [187] R. Gröber, A. Maier and T. Rauh, *Reconstruction of top-quark mass effects in Higgs pair production and other gluon-fusion processes*, *JHEP* **03** (2018) 020 [1709.07799].
- [188] R. Bonciani, G. Degrandi, P. P. Giardinò and R. Gröber, *Analytical Method for Next-to-Leading-Order QCD Corrections to Double-Higgs Production*, *Phys. Rev. Lett.* **121** (2018) 162003 [1806.11564].

Bibliography

- [189] J. Davies, G. Mishima, M. Steinhauser and D. Wellmann, *Double-Higgs boson production in the high-energy limit: planar master integrals*, *JHEP* **03** (2018) 048 [[1801.09696](#)].
- [190] M. Grazzini, G. Heinrich, S. Jones, S. Kallweit, M. Kerner, J. M. Lindert et al., *Higgs boson pair production at NNLO with top quark mass effects*, *JHEP* **05** (2018) 059 [[1803.02463](#)].
- [191] R. Grober, M. Muhlleitner, M. Spira and J. Streicher, *NLO QCD Corrections to Higgs Pair Production including Dimension-6 Operators*, *JHEP* **09** (2015) 092 [[1504.06577](#)].
- [192] R. Grober, M. Muhlleitner and M. Spira, *Higgs Pair Production at NLO QCD for CP-violating Higgs Sectors*, *Nucl. Phys.* **B925** (2017) 1 [[1705.05314](#)].
- [193] D. de Florian, I. Fabre and J. Mazzitelli, *Higgs boson pair production at NNLO in QCD including dimension 6 operators*, *JHEP* **10** (2017) 215 [[1704.05700](#)].
- [194] N. Deutschmann, C. Duhr, F. Maltoni and E. Vryonidou, *Gluon-fusion Higgs production in the Standard Model Effective Field Theory*, *JHEP* **12** (2017) 063 [[1708.00460](#)].
- [195] G. Cullen, N. Greiner, G. Heinrich, G. Luisoni, P. Mastrolia, G. Ossola et al., *Automated One-Loop Calculations with GoSam*, *Eur. Phys. J.* **C72** (2012) 1889 [[1111.2034](#)].
- [196] G. Cullen et al., *GOSAM-2.0: a tool for automated one-loop calculations within the Standard Model and beyond*, *Eur. Phys. J.* **C74** (2014) 3001 [[1404.7096](#)].
- [197] C. Degrande, C. Duhr, B. Fuks, D. Grellscheid, O. Mattelaer and T. Reiter, *UFO - The Universal FeynRules Output*, *Comput. Phys. Commun.* **183** (2012) 1201 [[1108.2040](#)].
- [198] A. Alloul, N. D. Christensen, C. Degrande, C. Duhr and B. Fuks, *FeynRules 2.0 - A complete toolbox for tree-level phenomenology*, *Comput. Phys. Commun.* **185** (2014) 2250 [[1310.1921](#)].
- [199] Z. Nagy, *Next-to-leading order calculation of three jet observables in hadron hadron collision*, *Phys. Rev.* **D68** (2003) 094002 [[hep-ph/0307268](#)].
- [200] G. P. Lepage, *VEGAS - an adaptive multi-dimensional integration program*, Tech. Rep. CLNS-447, Cornell Univ. Lab. Nucl. Stud., Ithaca, NY, Mar, 1980.
- [201] T. Hahn, *CUBA: A Library for multidimensional numerical integration*, *Comput. Phys. Commun.* **168** (2005) 78 [[hep-ph/0404043](#)].
- [202] A. D. Martin, W. J. Stirling, R. S. Thorne and G. Watt, *Parton distributions for the LHC*, *Eur. Phys. J.* **C63** (2009) 189 [[0901.0002](#)].

- [203] A. Buckley, J. Ferrando, S. Lloyd, K. Nordström, B. Page, M. Rüfenacht et al., *LHAPDF6: parton density access in the LHC precision era*, *Eur. Phys. J.* **C75** (2015) 132 [1412.7420].
- [204] ATLAS collaboration, *Measurements of the Higgs boson production and decay rates and coupling strengths using pp collision data at $\sqrt{s} = 7$ and 8 TeV in the ATLAS experiment*, *Eur. Phys. J.* **C76** (2016) 6 [1507.04548].
- [205] CMS collaboration, *Combined measurements of the Higgs boson’s couplings at $\sqrt{s} = 13$ TeV*, Tech. Rep. CMS-PAS-HIG-17-031, CERN, Geneva, 2018.
- [206] CMS collaboration, *Search for Higgs boson pair production in the $\gamma\gamma b\bar{b}$ final state in pp collisions at $\sqrt{s} = 13$ TeV*, *Phys. Lett.* **B788** (2019) 7 [1806.00408].
- [207] ATLAS collaboration, *Search for pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state using proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *JHEP* **01** (2019) 030 [1804.06174].
- [208] ATLAS collaboration, *Search for Higgs boson pair production in the $\gamma\gamma b\bar{b}$ final state with 13 TeV pp collision data collected by the ATLAS experiment*, *JHEP* **11** (2018) 040 [1807.04873].
- [209] C.-R. Chen and I. Low, *Double take on new physics in double Higgs boson production*, *Phys. Rev.* **D90** (2014) 013018 [1405.7040].
- [210] S. Dawson, A. Ismail and I. Low, *What’s in the loop? The anatomy of double Higgs production*, *Phys. Rev.* **D91** (2015) 115008 [1504.05596].
- [211] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel et al., *Scikit-learn: Machine learning in Python*, *Journal of Machine Learning Research* **12** (2011) 2825.
- [212] F. F. Freitas, C. K. Khosa and V. Sanz, *Exploring SMEFT in VH with Machine Learning*, 1902.05803.
- [213] CMS collaboration, *Measurement and interpretation of differential cross sections for Higgs boson production at $\sqrt{s} = 13$ TeV*, *Phys. Lett.* **B792** (2019) 369 [1812.06504].
- [214] F. Chollet et al., “Keras.” <https://keras.io>, 2015.
- [215] M. Abadi et al., “TensorFlow: Large-scale machine learning on heterogeneous systems.” <https://www.tensorflow.org>, 2015.
- [216] D. P. Kingma and J. Ba, *Adam: A Method for Stochastic Optimization*, 1412.6980.
- [217] CMS collaboration, *Measurement of inclusive and differential Higgs boson production cross sections in the diphoton decay channel in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **01** (2019) 183 [1807.03825].

Bibliography

- [218] ATLAS collaboration, *Combined measurements of Higgs boson production and decay using up to 80 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13\text{ TeV}$ collected with the ATLAS experiment*, *Phys. Rev.* **D101** (2020) 012002 [1909.02845].
- [219] M. A. Shifman, A. I. Vainshtein, M. B. Voloshin and V. I. Zakharov, *Low-Energy Theorems for Higgs Boson Couplings to Photons*, *Sov. J. Nucl. Phys.* **30** (1979) 711.
- [220] C. Grojean, E. Salvioni, M. Schlaffer and A. Weiler, *Very boosted Higgs in gluon fusion*, *JHEP* **05** (2014) 022 [1312.3317].
- [221] M. Schlaffer, M. Spannowsky, M. Takeuchi, A. Weiler and C. Wymant, *Boosted Higgs Shapes*, *Eur. Phys. J.* **C74** (2014) 3120 [1405.4295].
- [222] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040 [hep-ph/0409146].
- [223] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, *JHEP* **11** (2007) 070 [0709.2092].
- [224] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **06** (2010) 043 [1002.2581].
- [225] C. Anastasiou, C. Duhr, F. Dulat, E. Furlan, T. Gehrmann, F. Herzog et al., *High precision determination of the gluon fusion Higgs boson cross-section at the LHC*, *JHEP* **05** (2016) 058 [1602.00695].
- [226] R. K. Ellis, I. Hinchliffe, M. Soldate and J. J. van der Bij, *Higgs Decay to $\tau^+\tau^-$: A Possible Signature of Intermediate Mass Higgs Bosons at the SSC*, *Nucl. Phys.* **B297** (1988) 221.
- [227] U. Baur and E. W. N. Glover, *Higgs Boson Production at Large Transverse Momentum in Hadronic Collisions*, *Nucl. Phys.* **B339** (1990) 38.
- [228] V. Del Duca, W. Kilgore, C. Oleari, C. Schmidt and D. Zeppenfeld, *Higgs + 2 jets via gluon fusion*, *Phys. Rev. Lett.* **87** (2001) 122001 [hep-ph/0105129].
- [229] V. Del Duca, W. Kilgore, C. Oleari, C. Schmidt and D. Zeppenfeld, *Gluon fusion contributions to $H + 2\text{ jet}$ production*, *Nucl. Phys.* **B616** (2001) 367 [hep-ph/0108030].
- [230] F. Campanario and M. Kubocz, *Higgs boson production in association with three jets via gluon fusion at the LHC: Gluonic contributions*, *Phys. Rev.* **D88** (2013) 054021 [1306.1830].

- [231] N. Greiner, S. Höche, G. Luisoni, M. Schönherr and J.-C. Winter, *Full mass dependence in Higgs boson production in association with jets at the LHC and FCC*, *JHEP* **01** (2017) 091 [1608.01195].
- [232] R. Boughezal, F. Caola, K. Melnikov, F. Petriello and M. Schulze, *Higgs boson production in association with a jet at next-to-next-to-leading order in perturbative QCD*, *JHEP* **06** (2013) 072 [1302.6216].
- [233] X. Chen, T. Gehrmann, E. W. N. Glover and M. Jaquier, *Precise QCD predictions for the production of Higgs + jet final states*, *Phys. Lett.* **B740** (2015) 147 [1408.5325].
- [234] R. Boughezal, F. Caola, K. Melnikov, F. Petriello and M. Schulze, *Higgs boson production in association with a jet at next-to-next-to-leading order*, *Phys. Rev. Lett.* **115** (2015) 082003 [1504.07922].
- [235] R. Boughezal, C. Focke, W. Giele, X. Liu and F. Petriello, *Higgs boson production in association with a jet at NNLO using jetiness subtraction*, *Phys. Lett.* **B748** (2015) 5 [1505.03893].
- [236] R. V. Harlander, T. Neumann, K. J. Ozeren and M. Wiesemann, *Top-mass effects in differential Higgs production through gluon fusion at order α_s^4* , *JHEP* **08** (2012) 139 [1206.0157].
- [237] T. Neumann and M. Wiesemann, *Finite top-mass effects in gluon-induced Higgs production with a jet-veto at NNLO*, *JHEP* **11** (2014) 150 [1408.6836].
- [238] T. Neumann and C. Williams, *The Higgs boson at high p_T* , *Phys. Rev.* **D95** (2017) 014004 [1609.00367].
- [239] S. P. Jones, M. Kerner and G. Luisoni, *Next-to-Leading-Order QCD Corrections to Higgs Boson Plus Jet Production with Full Top-Quark Mass Dependence*, *Phys. Rev. Lett.* **120** (2018) 162001 [1802.00349].
- [240] R. Bonciani, V. Del Duca, H. Frellesvig, J. M. Henn, F. Moriello and V. A. Smirnov, *Two-loop planar master integrals for Higgs \rightarrow 3 partons with full heavy-quark mass dependence*, *JHEP* **12** (2016) 096 [1609.06685].
- [241] M. Becchetti, R. Bonciani, V. Casconi, V. Del Duca and F. Moriello, *Planar master integrals for the two-loop light-fermion electroweak corrections to Higgs plus jet production*, *JHEP* **12** (2018) 019 [1810.05138].
- [242] R. Bonciani, V. Del Duca, H. Frellesvig, J. M. Henn, M. Hidding, L. Maestri et al., *Evaluating two-loop non-planar master integrals for Higgs + jet production with full heavy-quark mass dependence*, 1907.13156.
- [243] H. Frellesvig, M. Hidding, L. Maestri, F. Moriello and G. Salvatori, *The complete set of two-loop master integrals for Higgs + jet production in QCD*, 1911.06308.

Bibliography

- [244] C. Anastasiou, C. Duhr, F. Dulat, E. Furlan, T. Gehrmann, F. Herzog et al., *CP-even scalar boson production via gluon fusion at the LHC*, *JHEP* **09** (2016) 037 [1605.05761].
- [245] R. V. Harlander, S. Liebler and H. Mantler, *SusHi Bento: Beyond NNLO and the heavy-top limit*, *Comput. Phys. Commun.* **212** (2017) 239 [1605.03190].
- [246] A. Azatov and A. Paul, *Probing Higgs couplings with high p_T Higgs production*, *JHEP* **01** (2014) 014 [1309.5273].
- [247] U. Langenegger, M. Spira and I. Strebel, *Testing the Higgs Boson Coupling to Gluons*, 1507.01373.
- [248] R. V. Harlander and T. Neumann, *Probing the nature of the Higgs-gluon coupling*, *Phys. Rev.* **D88** (2013) 074015 [1308.2225].
- [249] S. Dawson, I. M. Lewis and M. Zeng, *Effective field theory for Higgs boson plus jet production*, *Phys. Rev.* **D90** (2014) 093007 [1409.6299].
- [250] M. Grazzini, A. Ilnicka, M. Spira and M. Wiesemann, *Modeling BSM effects on the Higgs transverse-momentum spectrum in an EFT approach*, *JHEP* **03** (2017) 115 [1612.00283].
- [251] M. Grazzini, A. Ilnicka and M. Spira, *Higgs boson production at large transverse momentum within the SMEFT: analytical results*, *Eur. Phys. J.* **C78** (2018) 808 [1806.08832].
- [252] J. Kuipers, T. Ueda, J. Vermaseren and J. Vollinga, *FORM version 4.0*, *Comput. Phys. Commun.* **184** (2013) 1453 [1203.6543].
- [253] T. Peraro, *Ninja: Automated Integrand Reduction via Laurent Expansion for One-Loop Amplitudes*, *Comput. Phys. Commun.* **185** (2014) 2771 [1403.1229].
- [254] T. Binoth, J. P. Guillet, G. Heinrich, E. Pilon and T. Reiter, *Golem95: A Numerical program to calculate one-loop tensor integrals with up to six external legs*, *Comput. Phys. Commun.* **180** (2009) 2317 [0810.0992].
- [255] G. Cullen, J. P. Guillet, G. Heinrich, T. Kleinschmidt, E. Pilon, T. Reiter et al., *Golem95C: A library for one-loop integrals with complex masses*, *Comput. Phys. Commun.* **182** (2011) 2276 [1101.5595].
- [256] A. van Hameren, *OneLOop: For the evaluation of one-loop scalar functions*, *Comput. Phys. Commun.* **182** (2011) 2427 [1007.4716].
- [257] M. Boggia et al., *The HiggsTools handbook: a beginners guide to decoding the Higgs sector*, *J. Phys.* **G45** (2018) 065004 [1711.09875].
- [258] T. Gehrmann, M. Jaquier, E. W. N. Glover and A. Koukoutsakis, *Two-Loop QCD Corrections to the Helicity Amplitudes for $H \rightarrow 3$ partons*, *JHEP* **02** (2012) 056 [1112.3554].

- [259] MIT, “Keras2cpp.” <https://github.com/gosha20777/keras2cpp>, 2019.
- [260] S. Frixione, Z. Kunszt and A. Signer, *Three jet cross-sections to next-to-leading order*, *Nucl. Phys.* **B467** (1996) 399 [[hep-ph/9512328](#)].
- [261] M. Cacciari and G. P. Salam, *Dispelling the N^3 myth for the k_t jet-finder*, *Phys. Lett.* **B641** (2006) 57 [[hep-ph/0512210](#)].
- [262] M. Cacciari, G. P. Salam and G. Soyez, *The anti- k_T jet clustering algorithm*, *JHEP* **04** (2008) 063 [[0802.1189](#)].
- [263] M. Cacciari, G. P. Salam and G. Soyez, *FastJet User Manual*, *Eur. Phys.J.* **C72** (2012) 1896 [[1111.6097](#)].

Acknowledgments

I would like to thank the following people for their support during my Ph.D and for make it very enjoyable.

First my day-by-day supervisor Gudrun Heinrich my lead during the last three years. Thank you for giving me the opportunity of working in the phenomenology group at the Max Planck Institute, for your support along the path and for all the things you taught me.

Gerhard Buchalla and Wolfgang Hollik for useful discussion and suggestions about the thesis.

My project collaborators for the hard work which lead us to our publications: G. Buchalla, A. Celis, G. Heinrich, S. Jones, M. Kerner, L. Scyboz.

All the members of the phenomenology group past and present for all the moments spent together and useful discussions: H. Bahl, L. Chen, C. Dapla, A. Hala, T. Hahn, G. Heinrich, S. Hessenberger, S. Jahn, D. Lombardi, G. Luisoni, L. Maestri V. Papara, C. Pietsch, H. Pires, L. Scyboz, S. Zoia.

My officemates Joao Pires and Long Chen, it has been an honor to spend my time at the institute with you. Thanks for all the help you gave me and the funny moments we shared.

The friends I found in Munich: Elia Bertoldo, Dr. Victor Diaz, Ludovic Scyboz and Edoardo Vitagliano, you have been my family in the last three years. I will miss you a lot guys.

Nevertheless all my roommates and the people who stopped at least once by Connollystrasse to have an Italian dinner: G. Ceribella, G. D'Amico, N. Ferreiro, F. Guescini, J. Jimenez Pena, M. Mancuso, M. Predikaka, K. Novoselc, A. Wright.

All my Italian friends who make my life incredible: Alessio, Emanuele, Fabio, Gioele, Giorgio B., Giorgio N., Luca, Mariateresa, Manuel.

My girlfriend Ylenia for constantly standing by me, even though sometimes it is very hard, for brighting my days and for your love that is the most important thing to me. My life would be not the same without you, I love you.

My sister Valentina. You are my counselor, my anchor, the person who always save me from driving mental. I want you to know how lucky I am to have you in my life.

My parents Stefania and Vito, there are no words to say how thankful I am for all the things you have done for me. You taught me everything, you made me the man I am and if today I am here is only because of you.

My grandma Gina, you are always with me.

My last thank goes to Physics, you have been my greatest love for almost ten years, you changed my way to see the world, you changed me. For me now it is time to move on but you will be always in my heart.